Characterization of early and mature electrophysiological biomarkers of patients with Parkinson's disease

Christensen, Julie Anja Engelhard

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Technical University of Denmark

Julie Anja Engelhard Christensen

Characterization of early and mature electrophysiological biomarkers of patients with Parkinson's disease

PhD thesis, May 2015

DTU Electrical Engineering
Department of Electrical Engineering

Lundbeck
REGION H
Glostrup Hospital

PhD dissertation by:
Julie Anja Engelhard Christensen

Main supervisor:
Helge B. D. Sørensen, Associate Professor MSK, MScEE, PhD, Department of Electrical Engineering, Technical University of Denmark

Co-supervisors:
Poul Jennum, Professor, MD, PhD, Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Rigshospitalet, Glostrup
Søren Rahn Christensen, PhD, Clinical Pharmacology Scientist, Clinical Pharmacology, H. Lundbeck A/S
Lars Arvastson, PhD, Specialist, Biostatistics, H. Lundbeck A/S

Technical University of Denmark
Department of Electrical Engineering
Ørsteds Plads
Building 349
DK-2800 Kgs. Lyngby
Denmark
www.elektro.dtu.dk
Phone no.: (+45) 45 25 38 00

Project period:	September 1st, 2011 - May 3rd, 2015
Class:	Public
Edition:	1st Edition
Remarks:	This thesis is submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering at Technical University of Denmark.

Preface

The research providing the foundation of this PhD dissertation has been carried out at the Department of Biomedical Engineering at the Technical University of Denmark in close cooperation with the Danish Center for Sleep Medicine at the Department of Clinical Neurophysiology at Rigshospitalet, Glostrup and the pharmaceutical company H. Lundbeck A/S. The research constitutes a partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering.

The current PhD dissertation presents the primary research carried out over a period starting from September 2011 to May 2015, where other activities such as teaching, taking courses, participating in conferences, supervising students through Bachelor's and Master's projects and staying as a visiting student researcher at Stanford Center for Sleep Sciences and Medicine also have been completed.

The PhD dissertation consists of current summary report, four journal papers, two patent applications and two conference papers. All papers are accepted, and are included as appendixes to this thesis.

Julie Anja Engelhard Christensen

Acknowledgements

As an appreciation of generous help, guidance and contributions to this dissertation, I would like to express my thanks to a substantial group of people.

First of all, I would like to thank my supervisors Helge B. D. Sørensen from the Technical University of Denmark, Poul J. Jennum from the Danish Center for Sleep Medicine as well as Lars Arvastson and Søren R. Christensen from H. Lundbeck A/S for providing expert guidance and feedback throughout the project. I really appreciate the meetings we have had, both the small one-to-ones and the group meetings, where we have spend time discussing various aspects of the project. Your different academic backgrounds and personalities have in my opinion made the project exciting throughout the period and always with clues and ideas to pursue, which I am very grateful for.

Next, I would like to express my deepest appreciation to all the co-authors of my publications. Fruitful research is a teamwork, and this dissertation would not have been possible without your generous help and contributions to the different studies. I feel very pleased that I have had the chance to collaborate with so many talented researchers from various institutions and countries.

A special thanks goes to my colleagues and friends from the Biomedical Signal Processing Group; especially Lykke Kempfner, Dorthe Bodholt Saadi and Henriette Koch for their valuable feedback and support and always willingness to discuss various academic as well as personal matters. At rough times you have kept a positive and joyful working environment, both at DTU and at Stanford, which I will remember with a smile.

Lastly, I would like to thank my family for their support and understanding the last four years where the workload at times has been above normal.

Abstract

Neurodegenerative diseases (NDD) are highly disabling and severe diseases, and become more common with increasing age. As no cure exist and as the aging population increases, NDDs are considered to be one of the most serious health problems facing modern society. The most elusive goal in the field of NDD is to find a neuroprotective agent, and if such treatment becomes available, it is essential that the patients can be identified as early as possible. Parkinson's disease (PD) is the second most common NDD, and early disease identification is an active field of research as no reliable markers yet exist [83]. Sleep disturbances are common non-motor symptoms of PD, and strong findings associating a specific sleep disorder ("iRBD") to Parkinsonism suggest that sleep disturbances might precede the clinical diagnosis of PD. Analysis of sleep thus hold potential to serve as early disease identification, but as the current standard for sleep analysis relies on manual scorings guided by standards designed to fit healthy and normal sleep, manual sleep analysis of pathological sleep lacks substance.

This dissertation hypothesizes that automated sleep analysis can identify altered patterns of EEG and EOG in pathological sleep and may serve to reveal PD biomarkers. The aims of this dissertation was to: 1) Develop full data-driven sleep models based on EEG, EOG or both, that can describe sleep in detail and can be used in the analysis of normal as well as pathological sleep. 2) Extract appropriate features from the automated sleep models describing alterations in the sleep patterns of patients with PD or iRBD. 3) Identify changes of sleep spindles in the EEG of patients with PD by extracting features describing spindle morphology.

The results showed that patients with PD or iRBD reflect 1) altered eye movements during sleep, 2) altered amount and stability of data-determined stages linked to N3 and REM sleep, 3) more REM-NREM sleep transitions determined by a data-driven model, 4) decreased spindle density and 5) altered spindle morphology compared to non-NDD subjects.

In conclusion, this dissertation illustrates how appropriate biomedical signal processing can be used to reveal indicative alterations in the sleep EEG and EOG of patients with iRBD and PD. The automated methods developed analyze sleep in a robust and standardized way and can be supportive for sleep evaluation. Conclusively, this dissertation contributes to the field of early PD identification, but substantiates the claim that no known PD biomarker is reliable enough to stand alone.

Resumé

Neurodegenerative sygdomme (NDD) er yderst invaliderende og alvorlige lidelser, som bliver mere udbredt med alderen. Der findes ingen kur og da den aldrende befolkning stiger, anses NDDs for at være et af de mest alvorlige sundhedsproblemer i det moderne samfund. Et aktuelt mål indenfor NDD forskning er at udvikle et neurobeskyttende middel, og hvis et sådan middel bliver udviklet, er det essentielt at kunne identificere patienterne så tidligt som muligt. Parkinson's sygdom (PD) er den næstmest udbredte NDD, og da der ikke findes en pålidelig markør for sygdommen, omhandler meget forskning netop dette. Søvnforstyrrelser er almindelige blandt patienter med PD, og forskningsresultater viser at der er en stærk sammenhæng mellem en bestemt søvnsygdom ("iRBD") og PD, som antyder at søvnforstyrrelser går forud for PD symptomerne. Tidlig sygdomsindikation kan derfor potentielt findes ved søvnanalyser, men da nutidens analyser er baseret på standarder udviklet for normale søvnmønstre, er analyserne af patologisk søvn mangelfulde.

Denne afhandling fremsætter hypotesen om, at man ved at analysere søvn automatisk kan identificere forandringer i elektroencefalografi (EEG) og elektrooculografi (EOG) i patologisk søvn, og derved måske afsløre mulige biomarkører for PD. Formålene med denne afhandling var at 1) udvikle data-drevne søvnmodeller baseret på EEG og/eller EOG, der kan beskrive søvn i detaljer og derved supplere de manuelle analyser af normal og patologisk søvn; 2) udfra de data-drevne søvnmodeller at udtrække karakteristiske egenskaber der beskriver forandrede søvnmønstrer i patienter med iRBD og PD; 3) identificere ændringer i søvnspindler i EEG'et fra patienter med PD ved at udtrække mål for spindel morfologien.

Resultaterne viste at patienter med PD eller iRBD har 1) ændret øjenbevægelser under søvn, 2) ændret fordeling og stabilitet af automatisk fundne søvnstadier, der henholdsvis minder om N3 og REM søvn, 3) flere REM-NREM transitioner fundet ved en data-dreven model, 4) færre spindler og 5) ændret spindel morfologi sammenlignet med kontrolpersoner.

Denne afhandling illustrerer hvordan biomedicinsk signalbehandling kan anvendes til automatisk at identificere EEG og EOG ændringer under søvn hos patienter med iRBD eller PD. De udviklede automatiske metoder analyserer søvn på en robust og standardiseret måde og kan supplere nutidens søvnevaluering. Afslutningsvis, bidrager denne afhandling til forskningen indenfor tidlig PD identifikation, men konkluderer samtidigt at ingen kendt PD biomarkør er pålidelig nok til at stå alene.

Abbreviations

AASM	American Academy of Sleep Medicine
ACE	Addenbrooke's cognitive examination
AUC	Area under the ROC curve
BF	Basal forebrain
BG	Basal ganglia
BMI	Body mass index
BT	Time in bed
DE	Dream enactment
DR	Dorsal raphe nucleus
DWT	Discrete Wavelet transform
ECG	Electrocardiography
EEG	Electroencephalography
EMG	Electromyography
EMs	Eye movements
EOG	Electrooculography
FDR	False discovery rate
FFT	Fast Fourier transform
FN	False negative
FP	False positive
GC	Group consensus
H&Y	Hoehn and Yahr
IN	Cortical interneurons
iRBD	idiopathic REM sleep behavior disorder
LC	Locus coeruleus
LDA	Latent Dirichlet Allocation
LDT	Laterodorsal tegmental nucleus
LM	Leg movements
LPT	Lateral pontine tegmentum
MnPO	Median preoptic nucleus
MP	Matching Pursuit

MR	Median raphe nucleus
MSA	Multiple System Atrophy
MSLT	Multiple Sleep Latency test
N1	NREM sleep stage 1
N2	NREM sleep stage 2
N3	NREM sleep stage 3
NB	Naive Bayes
NDD	Neurodegenerative disorders
NREM	Non-REM
PC	Precoeruleus area
PD	Parkinson's disease
PLM	Periodic leg movements
PLMD	Periodic leg movement disorder
PPT	Pedunculopontine tegmental nucleus
PRE	Prethalamic fibers
PSG	Polysomnography
PY	Cortical pyramidal cells
RBD	REM sleep behavior disorder
RBDSQ	RBD screening questionnaire
RE	Thalamic reticular neurons
REM	Rapid eye movement
ROC	Receiver Operating Characteristic
RSWA	REM sleep without atonia
SC	Superior colliculus
SCRDA	Shrunken Centroids Regularized Discriminant Analysis
SE	Sleep efficiency
SL	Sleep latency
SLD	Sublaterodorsal region
SS	Sleep spindle
SVM	Support Vector Machine
TC	Thalamocortical relay cells
TMN	Tuberomammillary nucleus
TN	True negative
TP	True positive
UPDRS	Unified Parkinson's disease rating scale
vlPAG	Ventrolateral periaqueductal gray
VLPO	Ventrolateral preoptic nucleus
VSW	Vertex sharp waves
W	Wake

Table of Contents

1 Introduction

Parkinson's disease (PD) involves a progressive loss of structure or function of neurons in the brain area controlling voluntary movements. Almost 1% of the population over the age of 60 suffer from the disease, and the incidence rate grows with age [59]. No cure or neuroprotective agent yet exist, and due to the increasing aging population, neurodegenerative diseases (NDD), such as PD, are considered to be one of the most serious health problems facing modern society.

The clinical diagnosis of PD relies on motor symptoms, although it is widely known that there exist a pre-symptomatic interval where the pathological process has begun, but motor signs are absent. The pre-symptomatic interval can last several years, and less apparent symptoms such as cognitive decline, depression, behavioral changes, gastrointestinal and cardiovascular dysfunction, hyposmia or obesity exist, but are typically not observed or not linked to PD before the motor signs appear. Twenty years ago, it was discovered that a sleep disorder called idiopathic rapid eye movement (REM) sleep behavior disorder (iRBD) is closely related to Parkinsonism as it often precede the clinical diagnosis [71] [72] [70]. This group of patients therefore provide an opportunity for detecting biomarkers of PD before the symptoms onset, and its discovery lead to a huge new research area within the sleep community.

Sleep analysis relies on manual scoring, although it has been criticized for its oversimplification, high inter-scorer variability and large time consumption. Moreover, the standard for sleep scoring is designed to fit healthy and normal sleep, making the scoring of pathological sleep of dubious quality. This might explain why no clear disease marker has yet been identified in sleep. Abnormal high muscle activity during REM sleep has been indicated to be a PD biomarker [36] [73], but to strengthen the precision and give a more diverse picture of the individual disease progression, several indicative biomarkers are needed. Analysis of EEG has, to date, played a relatively minor role in this search, although it obviously hold the potential of contributing to early disease identification. This project proposes to use EEG and associated modalities such as EOG to characterize patients with PD or iRBD relative to control subjects. By employing appropriate biomedical signal processing algorithms, the characterization is expected to reveal new potential disease markers exempted from manual scorings that in future could result in higher precision of early disease identification that is possible today.

1.1 Problem statement

The most elusive goal in neurodegeneration including PD is the discovery of a neuroprotective agent that could slow down or stop the degenerative process. If a such treatment becomes available, it is essential that the patients can be identified as early as possible, also to let partially effective treatments have a much more powerful impact. This is the primary motivation for identifying markers for the pre-symptomatic interval of PD, and strong findings suggest that sleep analysis possess this potential. However, manual assessment of pathological sleep is problematic, as the sleep scoring standard do not guide the scoring of abnormal and altered sleep. This leads to imprecise analyses and may hinder identification of a robust PD biomarker.

1.2 Project objectives

With the motivation and problem statement discussed above, the overall objective of this PhD project was to provide new knowledge to the research of PD by analyzing sleep EEG and EOG patterns by use of appropriate biomedical signal processing algorithms. Specifically, the following focus points can be stated for this project:

- To develop a full data-driven and unsupervised model based on EOG, that can give a detailed description of the patterns of eye movements during sleep in non-NDD subjects as well as in patients with PD or iRBD.
- To develop a full data-driven and unsupervised model based on EEG, that can give a detailed description of the EEG patterns during sleep in non-NDD subjects as well as in patients with PD or iRBD.
- To develop a full data-driven and supervised model based on EEG and EOG, that can give a detailed description of sleep in non-NDD subjects as well as in patients with PD or iRBD.
- To extract appropriate features from the automated models describing deficits or alterations in the EOG, EEG and sleep patterns of patients with PD or iRBD compared to non-NDD subjects.
- To develop a database for sleep spindles in the EEG to be used when investigating spindle alterations in patients with PD.
- To identify changes of sleep spindles in the EEG of patients with PD by extracting features describing sleep spindle morphology.

1.3 Thesis outline and contributions

This dissertation is structured in three parts, each of them dealing with one specific area that could hold for potential biomarkers of PD. Besides the three chapters, it contains this introduction, preliminaries, and a conclusion where the findings are summarized and future aspects are stated and discussed.

Chapter 2 provides the basis of PD, sleep and an overview of the different research areas in the field of early identification of PD.

Chapter 3 presents our findings from working on eye movements during sleep. It encompasses results presented in two conference papers, paper I and II. It was decided to patent some of the work done on eye movements during sleep, which is presented in Patent I.

Chapter 4 presents our findings from working on a data-driven topic model used for sleep EEG and EOG analysis in PD. It presents the development of data-driven models based on EEG and EOG (in Paper III) as well as our findings when using such a model to characterize sleep stage transitions and stability in patients with PD (paper IV).

Chapter 5 presents our work done on an EEG sleep event called sleep spindles. It describes the development and utility of an automatic spindle detector (paper V), and our findings when characterizing spindle morphology in patients with PD (paper VI). It was decided to patent some of the work described in paper V which is presented in Patent II.

Chapter 6 summarizes our findings and suggestions for potential early PD biomarkers. It encompasses the final conclusions on the results of the project as well as ideas and directions for future work.

The scientific contributions of the PhD project are mainly covered by four journal papers, two conference papers, two patent applications and eight co-authored papers. These are listed as follows:

Patents:

- Application No.: PCT/EP2013/067297 / Publication No.: WO 2014/029764
  Applicant: Technical University of Denmark
  Title: Method for detection of an abnormal sleep pattern in a person
  Date of Priority: 20 Aug 2012
  Inventors: Christensen JAE, Sorensen HBD, Jennum P, Christensen SR, Arvastson L

- Application No.: 14/290,402 / Publication No.: US 2015/0080671
  Applicant: Technical University of Denmark
  Title: Sleep spindles as biomarker for early detection of neurodegenerative disorders
  Date of Priority: 29 May 2013
  Inventors: Christensen JAE, Kempfner L, Jennum P, Sorensen HBD, Arvastson L, Christensen SR

Journal papers:

- **Christensen JAE**, Jennum P, Koch H, Frandsen R, Zoetmulder M, Arvastson L, Christensen SR, Sorensen HBD. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clinical Neurophysiology 2015 [in press].

- **Christensen JAE**, Nikolic M, Warby SC, Koch H, Zoetmulder M, Frandsen R, Moghadam KK, Sorensen HBD, Mignot E, Jennum P. Sleep spindle alterations in patients with Parkinson's disease. Frontiers in Human Neuroscience 2015: 9:no 233.

- **Christensen JAE**, Zoetmuleder M, Koch H, Frandsen R, Arvastson L, Christensen SR, Jennum P, Sorensen HBD. Data-driven modeling of sleep EEG and EOG reveals characteristics indicative of pre-Parkinson and Parkinson's disease, Journal of Neuroscience Methods 2014; 235:262-76.

- **Christensen JAE**, Kempfner J, Zoetmulder M, Leonthin HL, Arvastson L, Christensen SR, Sorensen HBD, Jennum P. Decreased sleep spindle density in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease, Clinical Neurophysiology 2014; 125:512-9.

Conference papers:

- **Christensen JAE**, Koch H, Frandsen R, Kempfner J, Arvastson L, Christensen SR, Sorensen, HBS, Jennum P. Classification of iRBD and Parkinson's disease patients based on eye movements during sleep. 35th Annual International Conference of the IEEE EMBC; 2013:441-4.

- **Christensen JAE**, Frandsen R, Kempfner J, Arvastson L, Christensen SR, Jennum P, Sorensen HBS. Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel. 34th Annual International Conference of the IEEE EMBS; 2012:2941-4.

Co-authored papers:

- Koch H, **Christensen JAE**, Frandsen R, Zoetmulder M, Arvastson L, Christensen SR, Jennum P, Sorensen HBS. Automatic sleep classification using a data-driven topic model reveals latent sleep states. Journal of Neuroscience Methods 2014; 235:130-7.

- Krohne LK, Hansen RB, **Christensen JAE**, Sorensen HBS, Jennum P. Detection of K-complexes based on the wavelet transform. 36th Annual International Conference of the IEEE EMBC; 2014:5450-3.

- Koch H, **Christensen JAE**, Frandsen R, Arvastson L, Christensen SR, Sorensen HBS, Jennum P. Classification of iRBD and Parkinson's patients using a general data-driven sleep staging model built on EEG. 35th Annual International Conference of the IEEE EMBC; 2013: 5793-6.

- Kempfner J, Jennum P, Nikolic M, **Christensen JAE**, Sorensen HBD. Sleep phenomena as an early biomarker for Parkinsonism. 35th Annual Internatonal Conference of the IEEE EMBC; 2013:5773-6.

- Kempfner J, Jennum P, Sorensen HBS, **Christensen JAE**, Nikolic M. Automatic sleep staging: from young adults to elderly patients using multi-class support vector machine. 35th Annual International Conference of the IEEE EMBC; 2013:5777-80.

- Hansen IH, Marcussen M, **Christensen JAE**, Jennum P, Sorensen HBS. Detection of a sleep disorder predicting Parkinson's disease. 35th Annual International Conference of the IEEE EMBC; 2013:5793-6.

- Kempfner J, Jennum P, Nikolic M, **Christensen JAE**, Sorensen HBS. Automatic detection of REM sleep in subjects without atonia. 34th Annual International Conference of the IEEE EMBC; 2012:4242-5.

- Wendt SL, **Christensen JAE**, Kempfner J, Leonthin HL, Jennum P, Sorensen HBS. Validation of a novel sleep spindle detector with high performance during sleep in middle aged subjects. 34th Annual International Conference of the IEEE EMBC; 2012:4250-3.

2

Preliminaries

Objective *Parkinson's disease (PD) is the second most common neurodegenerative disorder (NDD) and has been estimated to affect 7-10 million people worldwide; a number that is increasing with the increasing aging population. Current treatment is purely symptomatic, and does not alter the underlying progress, but if a neuroprotective agent becomes available, intervention at an early stage is essential. During the last decades, it has become clear that patients with idiopathic rapid-eye-movement (REM) sleep behavior disorder (iRBD) are at high risk of developing PD, and therefore research has focused on sleep data in the search for PD biomarkers. Thus, this chapter will present the basic knowledge on PD and sleep, and will give an overview of the different research areas carried out in the search for PD biomarkers.*

2.1 Parkinson's disease

Neurodegenerative disorders (NDD) is an umbrella term for disorders were neurons are progressively degenerated. NDDs are among the most serious health problems facing modern society, especially because many of them become more common with advancing age. NDDs include the disorders Alzheimer's disease, Parkinson's disease (PD), Dementia, Amyotrophic lateral sclerosis and Huntington's disease. The characteristics and pathology of the different disorders depend on the function of the neurons affected, but due to huge inter-subject variability in symptom profile and disease progression, some of the diseases can in a clinical setting easily be confused with one another. There are many similarities between the different NDDs, an essential one being atypical protein assemblies, that may or may not be the main reason for the progressively worsening of the diseases. The treatment of all NDDs is purely symptomatic and do not cure nor alter the underlying process.

PD is the second-most common NDD after Alzheimer's disease and has been estimated to affect 0.5-1% among persons 65-69 years of age and 1-3% among persons 80 or more years of age [59]. Onset of PD is rarely before age 50 years, and the incidence rate increases with age with a sharp increase seen after age 60 years [23]. PD is characterized by Lewy body aggregations of a protein called alpha-synuclein [29]. The Lewy body aggregations typically starts in caudal areas of the brain and progress anteriorly [6]. When they reach the substantia nigra located in the midbrain, dopaminergic neurons die or fail to function resulting in a

reduction or depletion of dopamine [29]. The decline in dopamine production is causative for the typical motor symptoms of PD, as dopamine is an essential neutrotransmitter in the controlling of voluntary movements. As the disease progresses and dopamine production declines, patients are left unable to control movement normally.

Clinical assessment of PD relies on identification of characteristic motor abnormalities, and the diagnosis is typically set by stating the presence of at least two of the four motor symptoms: 1) resting tremor, 2) bradykinesia, 3) rigidity and 4) postural imbalance. Additionally, a lot of PD comorbidity exist including mental disorders, autonomic and gastrointestinal dysfunction and severe sleep disturbances, all of which considerable impair the quality of life of PD patients and relatives [49]. The non-dopaminergic and non-motor symptoms may appear years before the clinical diagnosis, and in advanced stages of PD, they can dominate the clinical picture and may be the most difficult symptoms to treat adequately [11]. This comes in line with the idea that the pathology of PD already has reached an advanced stage before the clinical appearance of PD [6]. Specifically, Braak et al. [6] has suggested that the neuronal damage in PD follows the Lewy-body distribution, which rise from the dorsal motor nucleus of the vague nerve in the medulla and in the olfactory bulb (stage 1). This stage is considered a pre-Parkinsonian state together with stage 2, where the inclusions have emerged through the subceruleus-ceruleus complex and the magnocellularis reticular nucleus. It is not considered a Parkinsonian state until involvement of the substantia nigra, the pedunculopontine nucleus and the amygdala (stage 3) as well as the temporal mesocortex (stage 4). Lastly, late-Parkinsonian states include initial involvement of the neocortex (stage 5) and at last an affection of nearly the entire neocortex (stage 6) [6].

Figure 2.1 schematically illustrates how PD progress according to Braak et al. [6]. The horizontal dotted line indicates the clinical manifestation stated as year 0. Above the line, the typical clinical symptoms are indicated together with a theoretical time line and the Hoehn and Yahr (H&Y) stages schematically stating the severity of the PD motor symptoms. Below the line are stated potential biomarkers of PD as these are the typical symptoms apparent before the clinical onset of PD. As seen in figure 2.1, a sleep disorder named rapid-eye-movement (REM) sleep behavior disorder (RBD) has been associated with PD. It has become increasingly clear that the presence of the idiopathic form of RBD (iRBD) even without the presence of motor or cognitive complaints confers a significant risk of conversion into PD or another synucleinopathy [38] [70] [72] [71] [11]. The discovery of iRBD provides opportunities to investigate trends of non-dopaminergic and non-motor signs of PD before the clinical onset, and several studies have focused on analysis of sleep data in the search for PD biomarkers [24] [11] [47] [46]. Additionally, early signs of sleep disturbances fit well in Braak's hypothesis, as the control of sleep is regulated by brainstem structures located close to where the Lewy body aggregation starts in early stages of PD [86].

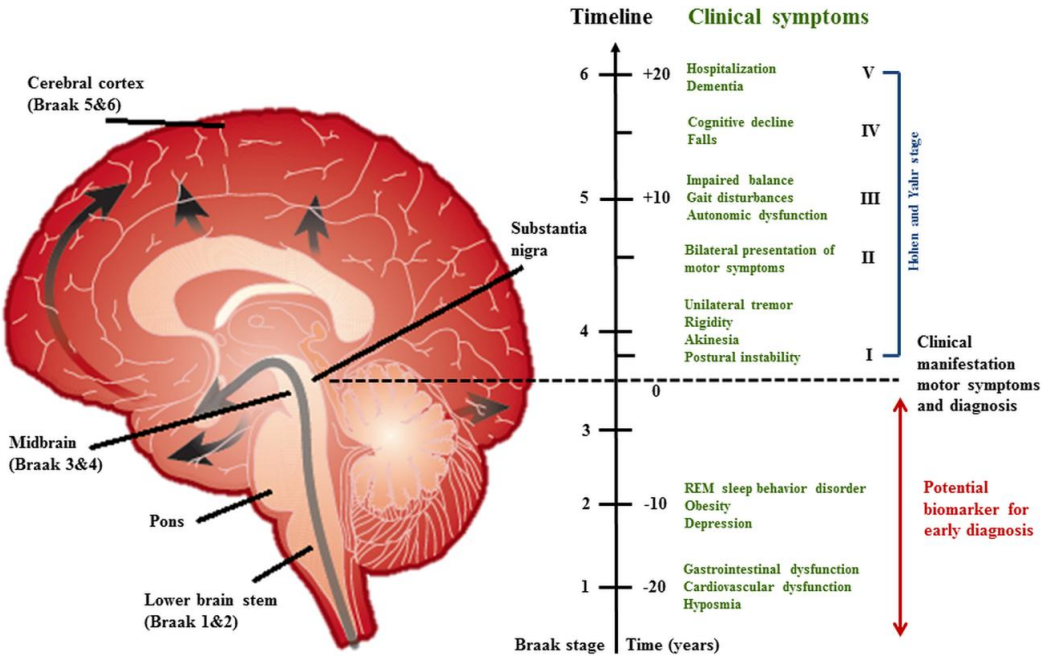


Figure 2.1: The progression of Lewy bodies according to Braak et al. [6]. Brain illustration taken from [1] and timeline taken from [82].

2.2 Sleep

During the last decades, sleep research has expanded tremendously, not only in the field of PD biomarkers, but also due to the increase in incidences of patients with apnea and other sleep disorders, and the fact that sleep has a major effect on one's normal well-being and functioning. Sleep is a vital and complex part of the human life and although we spent up to a third of our lifetime asleep, the functioning and mechanisms of sleep are still not fully understood. A lot of factors influence the sleep-wake cycle including body temperature, light exposure, hormone levels and others. Illness and stress can easily change the overall sleep structure from one night to another, and lack of sleep can influence cognitive functions. The overall sleep architecture changes with age, and the need and experience of sleep is very individual.

2.2.1 Sleep architecture and scoring

Analysis of sleep requires a sleep recording called polysomnography (PSG), which typically includes electroencephalography (EEG), electrooculography (EOG), electromyography (EMG), electrocardiography (ECG), respiratory airflow and peripheral pulse oximetry [34]. According to the newest standard for scoring sleep and associated events (described in the American Academy of Sleep Medicine (AASM) standard) [34], sleep annotation is done by manually assigning non-overlapping periods of 30 seconds to either wakefulness (W), REM sleep or one of three non-REM (NREM) sleep stages (N1-N3) dependent on the degree of

drowsiness. The scoring of sleep relies on identification of specific characteristics of the PSG signals including the clinical EEG frequency bands (delta, theta, alpha, beta), EEG microsleep events (sleep spindles (SS), K-complexes, sawtooth waves), EOG microsleep events (eye blinks, slow and rapid eye movements (EMs)) and EMG appearance and tonus [34]. These characteristics serve as hallmarks for the individual stages and can, based on their position within an epoch, define the initiation or termination of sleep stages.

Figure 2.2 provides an overview of the typical characteristics of each of the five sleep stages, and table 2.1 states the approximately stage distribution and the definitions of the microsleep events hallmarking the sleep stages.

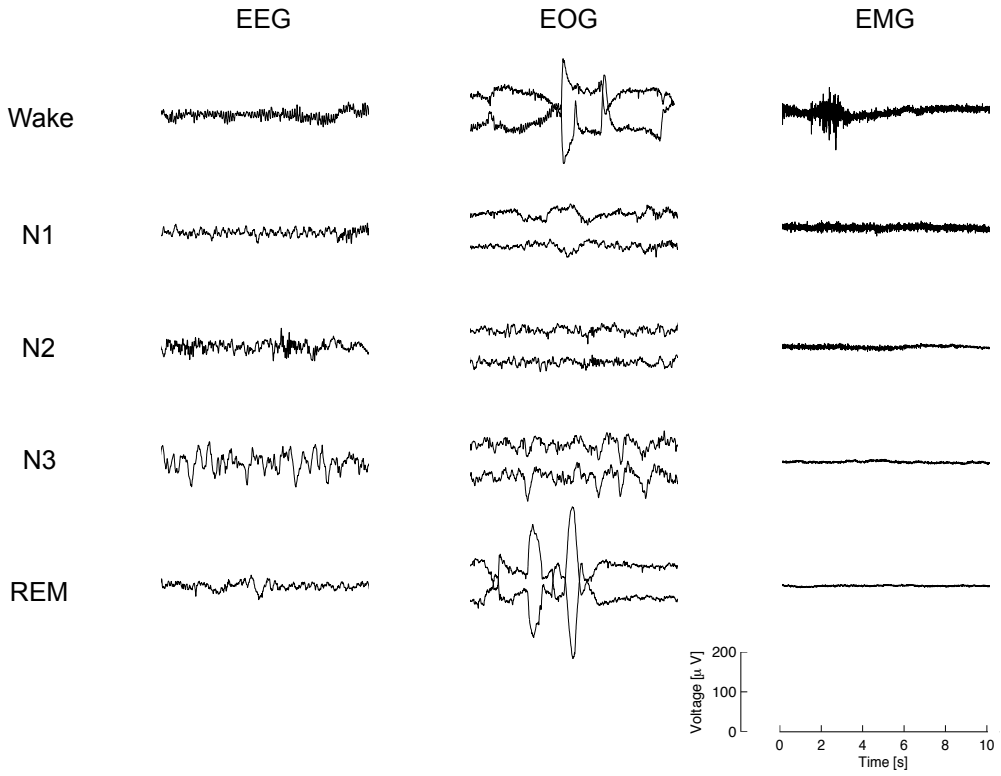


Figure 2.2: Polysomnographic characteristics of the different sleep stages.

During wakefulness, the EEG is typically low in amplitude and high in frequency (alpha (8-13 Hz) or beta (14-30 Hz) activity). When the eyes are closed, alpha activity dominates the picture and is mostly apparent at occipital derivations. When concentrating with open eyes, the EEG frequencies rise to beta activity with lower amplitudes, and when relaxing and becoming more drowsy, the EEG frequencies are slowing to become theta activity (4-7 Hz) with higher amplitudes.

Typically, N1 is the first sleep stage to enter from wakefulness. Here, theta activity dominates the picture and is mostly apparent at central and frontral derivations. Alpha activity is diminished and in some individuals vertex sharp waves (VSW) can be seen during N1. These are sharply contoured waves lasting less than half a second, and their appearance is one of

Stage	Percent of total time in bed	Definitions	
Wake	<5 %	EEG: EOG: EMG:	Alpha activity (8-13 Hz) when eyes closed Eye blinks, reading or rapid eye movements Variable amplitude, high tonus
N1	5-10 %	EEG: EOG: EMG:	Low amplitude, mixed frequency (predominantly 4-7 Hz) Vertex sharp waves: Sharply contoured, duration < 0.5 s Slow eye movements Variable amplitude, lower tonus than in wake
N2	45-55 %	EEG: EOG: EMG:	Low amplitude, mixed frequency activity K-complexes: Low frequency (0.5-2Hz), high amplitude, predominantly frontally Sleep spindles: Frequency of 11-16 Hz, duration of 0.5-3 s, predominantly centrally or frontally No or slow eye movements Variable amplitude, usually lower tonus than in N1
N3	15-20 %	EEG: EOG: EMG:	Large amplitude (>75 μV), slow wave activity (0.5-2 Hz) No eye movements, EEG activity might be reflected Usually lower tonus than in N2
REM	20-25 %	EEG: EOG: EMG:	Low amplitude, mixed frequency (predominantly 4-7 Hz), slow alpha activity in some individuals Sawtooth waves: Sharply contoured waves (2-6 Hz), predominantly centrally Rapid eye movements Lowest tonus

Table 2.1: Percent of total time in bed as well as EEG, EOG and EMG characteristics of the different sleep stages.

the hallmarks of sleep onset [34]. As K-complexes, VSW can be systematically excited by stimuli and is though to reflect an active inhibitory process designed to facilitate sleep onset [17].

N1 sleep is seen as a stage between drowsy wakefulness and the stable sleep stage N2. We spend approximately half our sleep time in N2 sleep which is characterized by an EEG activity of low amplitudes and mixed frequencies and presence of micro-sleep events such as SS and K-complexes. Essential restorative functions are performed during N2 sleep and maintaining sleep is therefore crucial. It is, however, also crucial for the brain to rouse in the face of danger. K-complexes are thought to suppress arousals in response to stimuli that are judged not to be dangerous [9]. They are often followed by a SS, which is thought to have a sleep-preserving role [30].

From N2, sleep usually continues into deep sleep, N3, which consists of slow-wave high-amplitude delta activity (0.5-2 Hz). The delta activity is dominantly at frontal derivations and is partly also reflected in the EOG. The SS density declines during N3, whereas the K-complex has been suggested to be a precursor to the delta waves produced during deep sleep [57] [22]. This stage takes up 15-20% of the total sleep times.

After approximately 90 minutes after sleep onset, REM sleep typically commences. Desynchronized and low-amplitude EEG activity dominates the picture, in some cases with presence of sharply contoured waves called sawtooth waves. Characteristic and name-given for this stage is the conjugated REMs seen in the EOG. It has been suggested that 80% of all dreaming occur during REM sleep, and a natural procession hereof is the typical atonia present during REM sleep. In healthy adults, the EMG tonus during REM sleep should be the lowest seen across all sleep stages, however, with small twitches that can break through the paralysis.

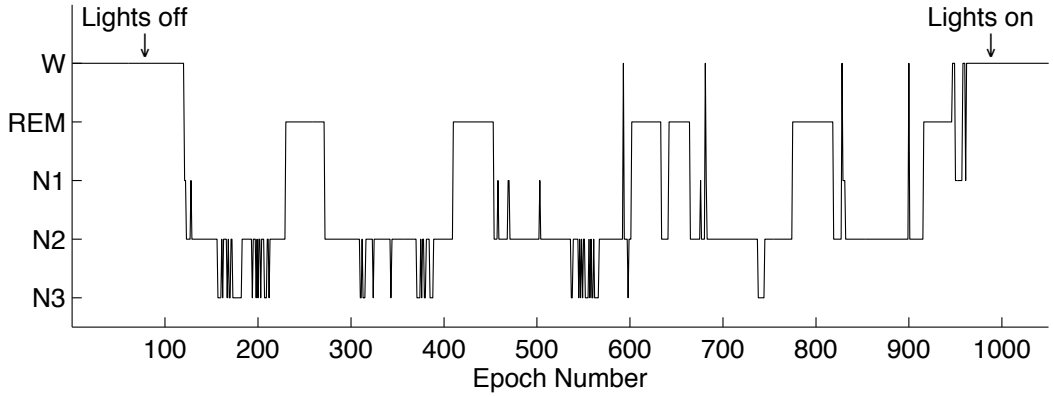


Figure 2.3: Manually scored hypnogram from a healthy male subject of age 54. Each epoch represents 30 seconds of sleep and the indications of lights off and lights on are self-reported timepoints.

Figure 2.3 shows an example of a hypnogram, which illustrates the temporal pattern of sleep summarizing the labeling of each 30-s epoch in a full night sleep. The shown example is a male of 54 years of age, with no known neurological disorders or sleep disturbances. As can be seen, a normal sleep typically starts with N1 followed by N2 and N3, thereafter turning into REM sleep. A normal healthy adult on average have four N1-N2-N3-REM sleep cycles during a full night sleep. The cycles gradually change throughout the night with most N3 sleep early in the night and more REM sleep close to waking up.

2.2.2 Changes in sleep with age

Age is the most consistent factor affecting sleep. With increasing age, alterations in the circadian and sleep homeostatic systems occur leading to poor sleep quality, increased sleep disturbances as well as difficulties in initiation and maintenance of sleep [55] [26]. Specifically, older adults have a sleep architecture that includes increased wake after sleep onset, increased time spent in N1 and N2, and reduced duration of N3 and REM sleep [26] [60]. Self-reports indicate increased time needed to fall asleep (i.e. increased sleep latency (SL)) and a slightly decrease in nighttime sleep amount, for which only the latter is confirmed by PSG data [26]. Whether sleep disturbances in older adults are due to changes in circadian rhythm directly or indirectly through e.g. changes in core body temperature and thereby melatonin secretion

is still unclear. Additionally, factors such as light exposure, menopause, medical conditions, medications, exercise and diet can effect melatonin production and may contribute to the sleep-wake disturbances in older adults [55].

The reduction in N3 sleep in elderly is profound, but might mainly reflect a decrease in the amplitude of delta activity rather than an absent of slow frequency activity [19]. K-complexes, possibly a precursor of delta waves, may reflect the ability of the brain to produce large amplitude synchronized waveforms, and with increasing age this micro-sleep structure is also found to be decreased in density and amplitude [19] [20]. Lastly, SS have been reported to be affected by age showing a reduction in density, amplitude and duration and are in general less well formed with increasing age [31] [20] [30] [44]. All of these age-related changes in the EEG is consistent with the neural degeneration seen with aging, and may indicate an inevitable age-related alteration of thalamo-cortical regulatory mechanisms.

2.2.3 Neuronal control of sleep

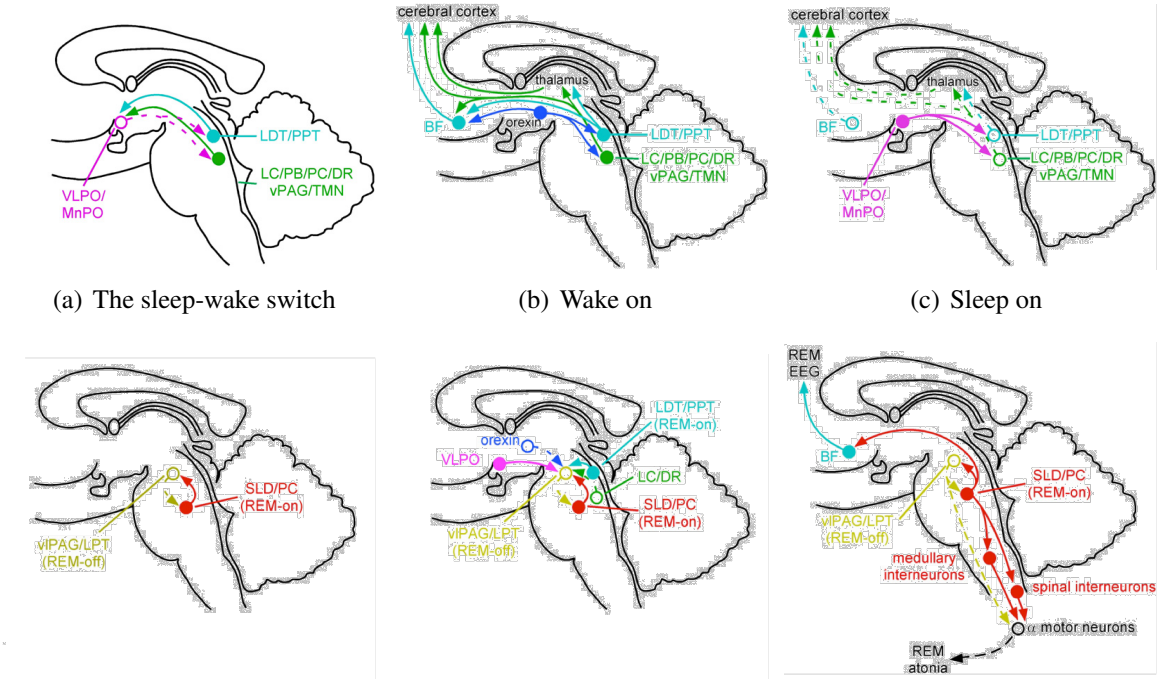
Sleep is strongly regulated by groups of neurons located in the brainstem and midbrain areas, which form reciprocal connections [51] [66] [67] [74]. These "sleep-wake switches" are mutually dependent and have been referred to as the wake-sleep and REM-NREM sleep switches, respectively.

Figure 2.4 shows illustrations of the sleep-wake switches and the neurons involved during wakefulness, NREM sleep and REM sleep as suggested by Saper et al. [67].

During wakefulness and REM sleep, cholinergic neurons located in the pedunculopontine and laterodorsal tegmental nucleus (PPT and LDT) fire rapidly, exciting ascending projections to the forebrain, notably the thalamus. This population of neurons forms one of two major branches involved in the ascending arousal system. The second branch consists of monoaminergic cell groups projecting to the lateral hypothalamus, basal forebrain (BF) and cerebral cortex [67] [74]. These include noradrenergic neurons of the locus coeruleus (LC), serotoninergic dorsal and median raphe nuclei (DR and MR), dopaminergic neurons adjacent to the DR, and histaminergic neurons of the tuberomammillary nucleus (TMN) [51] [67]. The neurons fire most rapidly during wakefulness, slowing down during NREM sleep and almost ceasing to fire during REM sleep.

The main sleep-promoting pathways consist of the ventrolateral and median preoptic nuclei (VLPO and MnPO), which act by inhibiting the circuits of the ascending reticular activating system. The mutually inhibitory relationship of the arousal- and sleep-promoting pathways together form the wake-sleep controlling mechanism, which generates complete transitions between waking and sleeping states [51] [67] [74].

The REM-NREM sleep transitions are controlled by the mutually inhibitory relationship of two populations of neurons located in the upper pons [67]. During wake and NREM sleep



(a) The sleep-wake switch (b) Wake on (c) Sleep on

(d) The REM-NREM sleep switch (e) REM on (f) REM on

Figure 2.4: Illustration of the neurons active during wake and sleep stages as suggested by Saper et al. [67]. (a): The basic neurons involved in the sleep-wake switch. (b): Wake-promoting projections from neurons in the upper brainstem provide input to the thalamus, hypothalamus, basal forebrain (BF) and cerebral forebrain. Neurons in the hypothalamus reinforce activity in the wake-promoting projections and directly excite the cerebral cortex and BF. (c): The sleep-promoting pathways inhibit the ascending arousal pathways in the hypothalamus and the brainstem. (d): The basic neurons involved in the REM-NREM switch. (e): Neurons in the locus coeruleus (LC) and dorsal raphe nucleus (DR) inhibit REM sleep by exciting REM-off neurons and inhibiting REM-on neurons, whereas neurons of the laterodorsal tegmental nucleus (LDT) promotes REM sleep by having opposite actions on these populations. Orexin neurons inhibit entry into REM sleep by exciting the REM-off neurons, whereas the ventrolateral preoptic nucleus (VLPO) promotes the entry into REM sleep by inhibiting this target. (f): Neurons in the sublaterodorsal region (SLD) activate inhibitory interneurons in the medulla and spinal cord leading to atonia. Projections from the parabrachial nucleus and precoeruleus area (PC) activates forebrain pathways leading to the characteristic EEG of REM sleep. All illustrations are taken from [67].

GABAergic (REM-off) neurons of the ventolateral periaqueductal gray (vlPAG) matter and the adjacent lateral pontine tegmentum (LPT) fire rapidly, inhibiting REM-on neurons in the sublaterodorsal region (SLD) [67]. The REM-on neurons are thought to be GABAergic or cholinergic by some groups [67], and glutamatergic by others [51] [50]. It has been proposed that the ascending branch of these neurons project into rostral brain areas responsible for the cortical activation during REM, while a descending branch projects into medullary glycinergic pre-motorneurons generating muscle atonia [10] [65]. Furthermore, Luppi et al. [51] [50] suggest that the descending branch also sends direct efferent projections to medullary GABA/glycinergic neurons, not only hyperpolarizing motorneurons but also dorsal horn sensory neurons involved in sensory processing. Most importantly, activity of the REM-on neurons inhibits the REM-off neurons and vice versa, which makes the REM-NREM controlling mechanism capable of generating complete transitions between REM and NREM sleep states.

Despite the mutually inhibitory loops involved in the two switching mechanisms, if either side of the two loops is weakened or injured, unwanted instability can occur in either of the states, irrespective of which side is damaged [74].

2.3 Early detection of Parkinson's disease

Looking at Braak's hypothesis of the progression path for the aggregations of Lewy bodies in synucleinopathies, it makes sense to search for PD biomarkers in physiological patterns linked to structures or neurons in the lower brainstem or the pons (see figure 2.1 and 2.5). As previously mentioned, an inevitable group of patients to analyze when searching for early PD biomarker is iRBD patients, as they possess a significant risk of conversion into synucleinopathies. As described by the AASM [2] in the 3rd edition of the international classification of sleep disorders, the diagnosis of RBD requires complaints or an anamnesis describing dream enactment (DE) behaviors as well as a manifestation of REM sleep without atonia (RSWA) as measured by PSG. This means that if either one of the two criteria is not obtained, e.g. no statement of DE or no REM sleep or RSWA in the recorded night, the diagnosis can not be stated. The idiopathic form of RBD is diagnosed when no concurrent neurological or other causing disease are found. This also entails that the idiopathic form of RBD only can be stated in cases where severe apnea is treated with CPAP.

The prevalence of RBD has previously been estimated to be 0.38% in elderly people [13] and 0.5% in the age group 15-100 years [61] [35], indicating this disease as rare among the general population. However, a recent study from Korea states the prevalence of RBD to be 2.01%, but the prevalence of subclinical RBD (i.e. increased muscle activity but with no history of DE behavior) to be 4.95% in elderly Koreans [40]. This suggests that more than twice as many people are affected by the disorder as first presumed, and these people may or may not have a risk just as high as the diagnosed patients for conversion into a synucleinopathy.

Stated in the diagnostic criteria, the most pronounced symptom of iRBD is the loss of atonia during REM sleep. A lot of research has been carried out in order to find reliable and objective measures of RSWA [41] [62] to be used to support diagnostic evaluations. This is an area associated with a lot of challenges, mainly because there is no exact definition of when atonia is present and when it is absent. AASM states that the REM sleep is associated with low EMG tonus, and in periods where REMs are absent, the scoring of sleep relies on EMG and EOG, which in cases of iRBD with RSWA, makes the manually scoring of REM sleep very difficult. With no clear REMs in the EOG and no clear EEG hallmarks of REM sleep, the scorer is left to decide whether the person is in REM sleep but with a high EMG tonus, has awakened or transitioned into another sleep stage as the criteria for scoring REM sleep are no longer met. Conclusively, the diagnosis of iRBD is associated with a lot of subjectivity, both from the patient in terms of stating DE, but also from the evaluator in terms of stating the presence of RSWA.

Besides analysis of RSWA, other sleep research has been carried out in iRBD patients, including analysis of slow wave sleep [47] [53], heart rate variability [75] [76] and EEG spectral power analysis [25] [68] [37]. However, quantitative sleep research in iRBD patients is still a new area and requires further efforts to clarify what abnormalities iRBD patients possess besides RSWA, and which of these present a true early marker of subsequent conversion into PD. The aim of this PhD project was to search for early PD biomarkers in electrophysiological signals, mainly EEG and EOG, recorded during sleep.

In figure 2.5 is given examples of potential PD biomarkers found during sleep and their neuroanatomic correlates. In this thesis, the focus has been on EMs, sleep stage switching and stability as well as characterization of SS in the EEG. Specifically, quantification of EMs during sleep and sleep stability are new research areas and not confirmed by any other group.

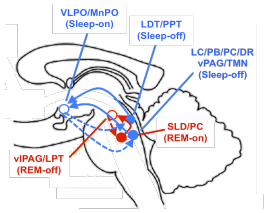
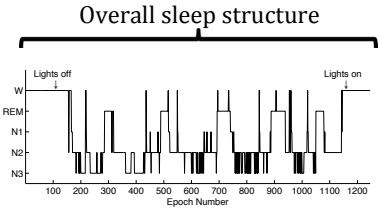
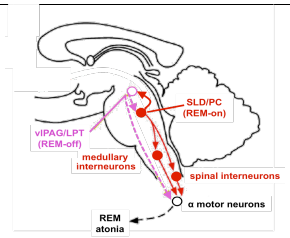
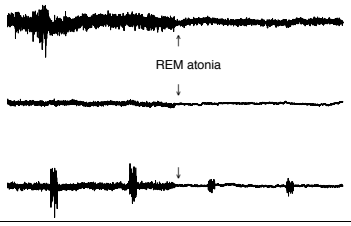
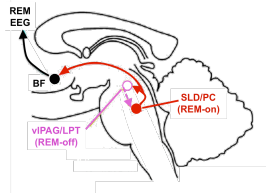
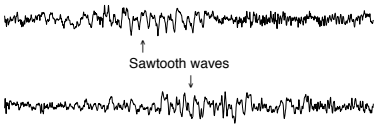
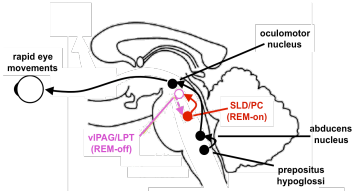
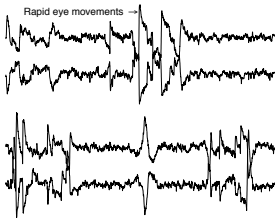
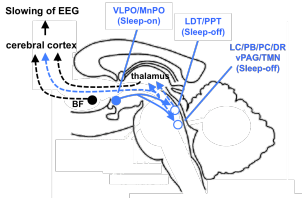
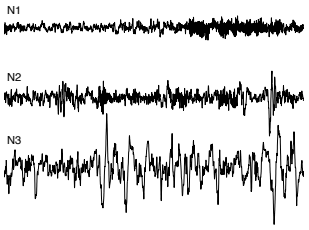
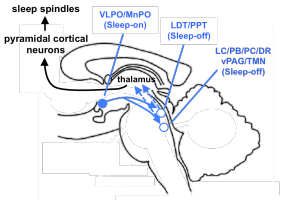
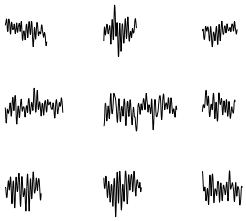
Neurons involved	Electrophysiological patterns	Signals investigated
		Hypnogram EEG: All derivations EOG: Left and right EMG: Chin, TIBR, TIBL
		EMG: Chin, TIBR, TIBL EOG: Left and right
		EEG: All derivations
		EOG: Left and right
		EEG: All derivations
		EEG: Central and frontal

Figure 2.5: Examples of potential early PD biomarkers found during sleep and their neuroanatomic correlates. The most investigated of these areas is analysis of the loss of atonia during REM sleep. Also, a lot of research has been carried out in EEG during slow wave sleep and REM sleep. This thesis focus on sleep spindles, eye movements and sleep stability, whereof in specific the two last-mentioned areas are new and not confirmed by any other group. Brain template taken from [67].

3

Eye movements during sleep

Objective *Eye movements (EMs) are controlled by neurons located in the lower brainstem, midbrain and frontal areas. Rapid EMs are present during REM sleep, indicating that the oculomotor muscles are not affected by the atonia during REM sleep. As the neurodegeneration is thought to start at lower brainstem areas, EMs during sleep might be affected in both iRBD and PD patients. Several studies have reported impairment of the oculomotor function in patients with PD during wakefulness, but no studies have investigated how neurodegeneration effect EMs during sleep. This chapter comprises two studies concerning this area. The first study is a pilot study investigating if features extracted from one EOG derivation can separate iRBD and PD patients from control subjects. As the first study revealed that features reflecting EMs were indicative of neurodegeneration, the second study focuses on EMs and tries to look deeper into the timely distribution of EMs by developing a data-driven model.*

3.1 Background

A saccade is a voluntary quick, simultaneous movement of both eyes. A lot of neurons and brain areas are involved in controlling a saccade, and the system can be divided in three major levels. The lowest level concerns the oculomotor muscles moving the eye balls, the intermediate level concerns the mechanisms for locating object and directing the eye movements (EMs), and the highest level concerns the decision-making of whether of not to move the eyes [8]. In figure 3.1 is seen a schematic representation of the main features of the saccadic control system.

The movement of the eyeballs themselves is controlled by neurons in the brainstem area. This movement has to be extremely fast, as eyes in motion are effectively blind. At the lowest level of the saccadic control system, neurons called omnipausers fire continuously and rapidly during rest. During a saccade, they pause for exactly the duration of the movement, regardless of its direction. The omnipausers control the firing of burst units, that fire rapidly during a saccade reflecting the direction of the saccade. The burst units control the tonic units, that fire tonically at one steady level before the saccade and at another steady level after the saccade reflecting the position of the eye [8]. The nucleus prepositus hypoglossi is thought to be the neural region that integrates the burst and tonic units. Controlling oculomotor neurons

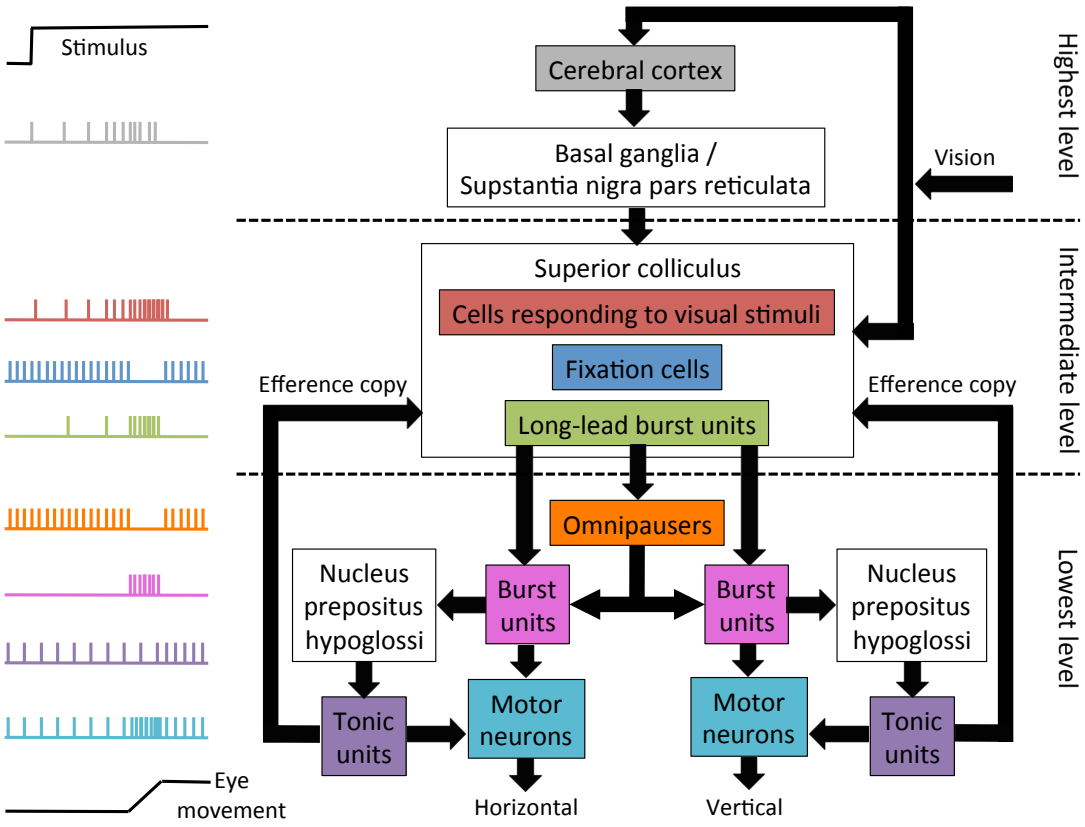


Figure 3.1: Schematic representation of the saccadic control system. Visual information is sent to the cortical areas for visual recognition and to the superior colliculus for target localization. When it is decided to look at a target, the basal ganglia and the substantia nigra pars reticulata lift their inhibitory blanket allowing the superior colliculus to initiate the processes needed to activate the oculomotor muscles. A schematically illustration of the activity patterns of the different neurons is seen to the left. Illustration modified from [8].

by these two signals ensures that saccades can happen as quick as a little more than 20 milliseconds.

At the intermediate level, objects are located and eyes are directed to the appropriate position. This is thought to be controlled by the superior colliculus (SC) located at the top of the brainstem. SC contain neurons that respond to visual stimuli at different places in the visual field as well as motor neurons that can direct and fixate the eyes. A model proposes that when a visual stimulus is registered at the SC from the retina fibers, neurons called long-lead bursters are stimulated at a site corresponding to a particular visual location that will lead the eyes to the appropriate spot in the outside world. Additionally, the visual stimulus will inhibit neurons that fire rapidly during fixation of the eyes. An inhibition of the fixation neurons lead to a pause in the firing from omnipausers which initiates the saccade itself. During the saccade, information from the oculomotor neurons is send back to the SC (efference copy), that continuously guide the gaze towards the fixation region. When this region is reached,

the omnipausers and the fixation neurons are turned back on again, and the eyes fixate. The fact that the oculomotor muscles do not carry any external load makes the oculomotor system very fast and precise, as it can work out where the eyes are positioned at any given time only by monitoring the signals being send to the muscles (efference copy).

The highest level of the saccadic control system concerns the decision-making, and involves several regions of the cerebral cortex [8]. Before the decision is made, the eyes are kept in position, which is achieved by continuously inhibiting the SC. The responsible tonically active inhibitory fibers project down from the substantia nigra pars reticulata in the basal ganglia. Fibers from the retina project to the visual cortex, where the visual image is broken down into its fundamental shapes used for visual recognition. The visual recognition occurs in several areas of the cerebral cortex, and when it is decided to look at a target, the blanket of inhibition is lifted briefly and locally, permitting the SC to generate an appropriate saccade [8].

3.1.1 Electrooculography

During a polysomnography (PSG), the recording of EMs is done by placing electrodes besides each outer canthus. As the eyeball acts as a dipole with a potential difference between the negative retina (posterior point of the eyeball) and the positive cornea (anterior point of the eyeball), the EOG electrodes will register negative and positive potentials as the eyes are moving. The simultaneously movement of the eyeballs entails the two EOG signals to be synchronized and anti-correlated. The EOG electrode nearest the retina will reflect a negative potential and the one nearest the cornea will reflect a positive potential.

The horizontal and vertical distance from the outer canthus to the electrode as well as the angle at which the eyes are gazing influence the relative amplitude change in the EOG signals. Also, during a PSG, the EMs are recorded by placing EOG electrodes a bit higher (right eye) or a bit lower (left eye) than the mid-line of the eyes. Therefore, the EOG signals recorded as part of a PSG typically do not show the same relative amplitude change as the eyes moves. Optimally, electrodes should be places above, under, and on either side of both eyes in order to fully reflect the vertical and horizontal EMs. During a PSG, however, the EOG is mostly used to reflect the presence of slow and rapid EMs.

3.1.2 Research hypothesis

The main hypothesis is that patients with iRBD and especially patients with PD reflect abnormal EMs during sleep. In iRBD patients, the neurodegeneration is thought to effect the lower level of the saccadic control system, leading to abnormal form of saccades. As the disease progresses, the intermediate and higher level might be affected as well, maybe also effecting the sustained inhibition of the SC leading to uncontrolled EMs in patients with PD.

The aim with this research area was to:

- reveal with a pilot study whether or not EMs during sleep have potential to be a biomarker for PD.
- develop a data-driven model based on data describing EMs during sleep, and investigate its usability to analyze and automatic classify control subjects and patients suffering from iRBD or PD.

3.2 Paper I: Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel

This study is a pilot study investigating whether EOG hold information that can be used to separate control subjects from iRBD and PD patients. The features are derived so they are easy to interpret to make it accessible to reveal whether indicative features reflect EMs or other physiological signals recorded at the EOG side.

The overall methodology of this study is seen in figure 3.2.

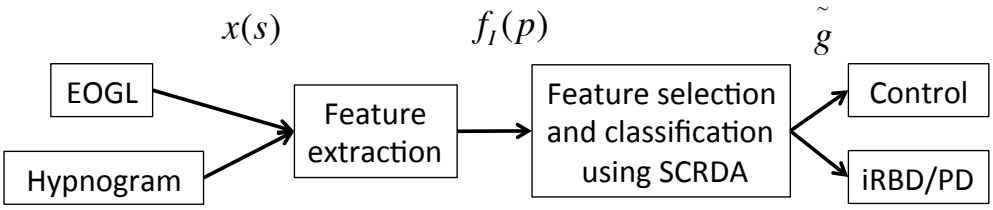


Figure 3.2: Overall methodology of study I concerning eye movements during sleep. The features $f_I(p)$ were extracted from each sleep epoch between lights off and lights on determined by the manually scored hypnogram. The Shrunken Centroids Regularized Discriminant Analysis (SCRDA) method was used for feature selection and classification. SCRDA outputs the class label $\tilde{g}$ for each subject classifying them as "control" or "patient".

3.2.1 Methods: Preliminary analysis of EOG features by the Shrunken Centroids Regularized Discriminant Analysis

Subjects and recordings

A total of ten control subjects, ten patients with iRBD and ten patients with PD were enrolled in this study. The evaluation of the patients included a comprehensive medical and medication history and a PSG analyzed according to the American Academy of Sleep Medicine (AASM)

standard [34]. If narcolepsy was suspected, a multiple sleep latency test (MSLT) was performed as well. Subjects taking anti-depressants, including hypnotics or medication known to affect sleep were excluded, though dopaminergic treatments were allowed and continued. The control subjects had no history of dream-enacting behavior, movement disorder or sleep disorders. Demographic data for the three groups analyzed is seen in table 3.1. The total number of epochs in between light off and light on is also provided.

	Controls	iRBD	PD
Total number	10	10	10
Male / Female	5 / 5	8 / 2	6 / 4
Age [years]	59.8 $\pm$ 8.4	59.0 $\pm$ 14.2	63.2 $\pm$ 8.4
Wake (%)	1173 (12)	1881 (18)	1882 (19)
REM (%)	2000 (21)	1731 (16)	1531 (15)
N1 (%)	678 (7)	1081 (10)	1275 (13)
N2 (%)	4443 (46)	4881 (46)	4073 (42)
N3 (%)	1347 (14)	1114 (10)	1084 (11)

Table 3.1: Demographic data and total number of epochs across the sleep stages between lights off and lights on for the three groups analyzed in study I.

The quality of the PSG recordings was individually evaluated and recordings with discontinued or artifact-contaminated channels were excluded. Besides the manually scored hypnogram, only the left side EOG channel was used in this study. The sampling frequency was 256 Hz.

Feature extraction

After removal of the 50 Hz powerline noise, each sleep epoch (N1, N2, N3 or REM determined by the manually scored hypnogram) was decomposed using the Discrete Wavelet Transform (DWT). In DWT, an input signal $x(n)$ is passed through a series of filters splitting the signal into its high and low frequency components, denotes as detail (D) and approximation (A) components, respectively. After filtration, the components contain redundancies and a downsampling by a factor of two is applied. The decomposition in DWT is based on shifting and scaling a mother wavelet finite in length. A mother wavelet that is contracted twice will have a twice higher central frequency and a twice smaller time duration, which entails a decrease in frequency resolution and an increase in time resolution with increasing frequencies [4]. By using DWT in this study, we obtained a high resolution at lower frequencies compared to higher frequencies, which was preferred as the low frequency ranges are the ones reflecting EMs.

For each sleep epoch, a DWT decomposition was carried out to level eight using the Daubechies 4 (db4) as the mother wavelet. The energy percentage and the common logarithm of the summed absolute signal values of the reconstructed detail subbands d2-d8 were computed, and a single feature vector, $f_I(p)$ was obtained for each subject p by taking the mean

(μ) and standard deviation (σ) of these values across all sleep epochs. For each subject, $f_I(p)$ hold 28 feature values described as,

$$f_I(p) = \begin{bmatrix} \left[f_{\mu(\%E)}\right]_{d2}^{d8} \\ \left[f_{\sigma(\%E)}\right]_{d2}^{d8} \\ \left[f_{\mu(\log_{10}E)}\right]_{d2}^{d8} \\ \left[f_{\sigma(\log_{10}E)}\right]_{d2}^{d8} \end{bmatrix} \quad \text{where} \quad [f_E]_{d2}^{d8} = \begin{bmatrix} f_{E_{d2}} \\ \vdots \\ f_{E_{d8}} \end{bmatrix} \tag{3.1}$$

and $\%E_{dx}$ and $\log_{10}E_{dx}$ indicate the energy percentage and the common logarithm of the summed absolute signal values of the reconstructed detail subbands dx, respectively. The features are summarized in table 3.2, and they reflect the energy distribution of EOG across different frequency ranges. As it is assumed that controls have periods with no EMs (low energy) and periods with rapid or slow EMs (higher energy), the standard deviation across the night might be different in between patients and controls.

Feature name	Feature explanation	Total no per subject
$\left[f_{\mu(\%E)}\right]_{d2}^{d8}$	Mean across all sleep epochs of the energy percentages in the reconstructed subbands $d2$-$d8$	7
$\left[f_{\sigma(\%E)}\right]_{d2}^{d8}$	Standard deviation across all sleep epochs of the energy percentages in the reconstructed subbands $d2$-$d8$	7
$\left[f_{\mu(\log_{10}E)}\right]_{d2}^{d8}$	Mean across all sleep epochs of the common logarithm of the summed absolute signal values of the reconstructed subbands $d2$-$d8$	7
$\left[f_{\sigma(\log_{10}E)}\right]_{d2}^{d8}$	Standard deviation across all sleep epochs of the common logarithm of the summed absolute signal values of the reconstructed subbands $d2$-$d8$	7

Table 3.2: Overview of the features computed in study I. For each sleep epoch, the energy percentage and the common logarithm of the summed absolute signal values were computed. The features reflect the mean or the standard deviation of these energy measures across all sleep epochs.

Feature selection and classification

To avoid overfitting, only a subset of features was used to classify the subjects. In this study, the Shrunken Centroids Regularized Discriminant Analysis (SCRDA) method that generalizes the idea of Nearest Shrunken Centroids into the classical discriminant analysis was used to evaluate the features and classify the subjects. This method was chosen as it is designed for classification problems where the number of features is larger or nearly the same as the number of observations, as in this case. Also, SCRDA is very suitable for feature eliminating purposes, enabling sparse models less likely to overfit. The criteria in the standard Linear Discriminant Analysis can be stated as:

$$\begin{gathered} x_{g,i} \in \text{population}\left(\tilde{g} = \arg\max_{g'} d_{g'}(x_{g,i})\right) \quad \text{where} \\ d_g(x) = x^T\Sigma^{-1}\mu_g - \frac{1}{2}\mu_g^T\Sigma^{-1}\mu_g + \log(\pi_g) \end{gathered} \tag{3.2}$$

where the observation $x_{g,i}$ is classified to the population $\tilde{g}$ that has the largest posterior probability of the observation when the prior probability π_g is included. Here, $\mu_g = \frac{1}{n_g}\sum_{i=1}^{n_g} x_{g,i}$ and $\Sigma = \frac{1}{2}\left(X - \overline{X}\right)\left(X - \overline{X}\right)^T$ describes the mean vector and the common covariance matrix of population g, respectively. X and $\overline{X}$ are $p \times n$ matrices holding the observations and the mean vectors column wise, respectively, and n_g is the number of samples in population g. In cases where the number of samples is small compared to the number of features, the covariance matrix is poorly estimated, and in SCRDA this singularity issue is managed by estimating the covariance matrix by use of a regularization parameter $0 \leq \alpha \leq 1$ [32], yielding the estimate

$$\widetilde{\Sigma} = \alpha\Sigma + (1-\alpha)D \quad \text{with} \quad D = \text{diag}(\Sigma). \tag{3.3}$$

When α is shifted toward zero, $\widetilde{\Sigma}$ tends to be the diagonal estimate of the covariance matrix D and when α is close to 1, the standard estimate of the covariance matrix Σ is chosen as the optimal choice. By changing α, SCRDA thereby perform well on all conditions.

Another useful aspect of the SCDRA is that it gains sparsity meaning that it removes variables that are thought to be noisy leaving only a small subset of variables for further investigation. This is gained by substitution the class means μ_g with the shrunken centroids μ_{gs}, which is achieved by first transforming μ_g to μ_{gs}^* and then shrinking it toward 0 by use of a tuning parameter Δ:

$$\mu_{gs}^* = \text{sign}\left(\widetilde{\Sigma}^{-1}\mu_g\right)\left(|\widetilde{\Sigma}^{-1}\mu_g| - \Delta\right)_+. \tag{3.4}$$

The subscript plus means positive part and entails $t_+ = t$ if $t > 0$ and zero otherwise. Finally, μ_{gs}^* is transformed back to get the shrunken centroid $\mu_{gs} = \widetilde{\Sigma}\mu_{gs}^*$. The shrunken centroids has the benefit of eliminating noise variable which does not contribute to classification. It automatically select the variables i.e. by eliminating a variable if the class centroid is zero for all classes [12] [32].

Conclusively, the criteria the SCRDA can be stated as:

$$\begin{gathered} x_{g,i} \in \text{population}\left(\tilde{g} = \arg\max_{g'} \tilde{d}_{g'}\left(x_{g,i}\right)\right) \quad \text{where} \\ \tilde{d}_g(x) = x^T\widetilde{\Sigma}^{-1}\tilde{\mu}_{gs} - \frac{1}{2}\tilde{\mu}_{gs}^T\widetilde{\Sigma}^{-1}\tilde{\mu}_{gs} + \log(\pi_g). \end{gathered} \tag{3.5}$$

Each subject was classified by a leave-one-out approach where the values of Δ and α were found by a 10-fold cross-validation on the 29 subjects not held out. The leave-one-subject out approach was chosen due to the limited number of subjects included in this project, and the 10-fold cross validation was chosen to avoid overfitting. The final values were found as the mean across the 30 runs, indirectly giving the optimal subset of features across the 30 runs.

3.2.2 Results

Table 3.3 summarizes the optimal subset of features found by the SCDRA method where the parameters Δ and α are the mean values across the 30 runs. It is seen that two of the features include frequencies in the range 1-2 Hz, which could reflect EMs. Feature number 1 includes frequencies in the range 32-64 Hz, which must reflect electromyographic (EMG) activity recorded at the EOG side. Feature number 2 includes frequencies in the range 16-32 Hz, which could reflect EMG activity and to a lesser degree also EEG activity recorded at the EOG side.

Original feature no.	Frequency range [Hz]	Description of feature
1	32-64	Mean of the logarithmic "energy" in d2
2	16-32	Mean of the logarithmic "energy" in d3
20	1-2	Mean of the percentage energy in d7
27	1-2	Standard deviation of the percentage energy in d7

Table 3.3: The optimal subset of features found by using the SCRDA method to classify controls and iRBD/PD patients in study I.

In figure 3.3 is seen the posterior probability of belonging to the NDD class using the leave-one-out SCRDA classification approach. It should be emphasized, that because an optimal set of parameters, Δ and α, was found for each subject, the included features for classifying each subject may vary. The four features presented in table 3.3 reflect the overall best subset. It was, however, noted that feature number 1 and 27 were chosen in all 30 runs indicating that these two might be the most discriminative ones.

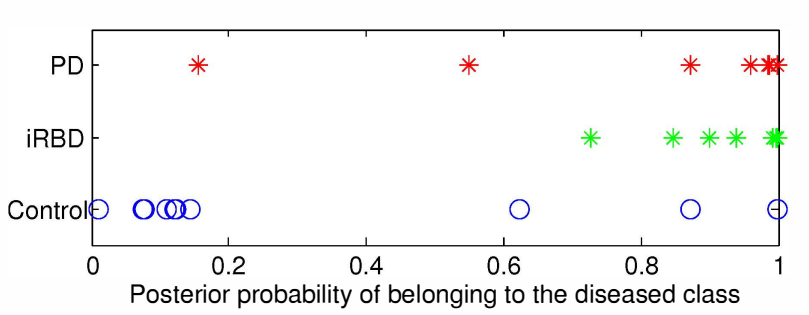


Figure 3.3: The posterior probability of belonging to the NDD class holding the iRBD and PD patients indicated as green and red stars, respectively. Controls are indicated as blue open circles. For each subject, the posterior probability was computed based on the leave-one-subject out classification approach by use of the discriminant function in equation 3.5, where the optimal values for Δ and α were obtained from a 10-fold cross-validation on the 29 subjects not held out. The closer to 1 (= 100%) a subject is, the higher posterior probability of belonging to the NDD class.

Following the criteria in SCRDA, it is seen that three control subjects were misclassified as patients (posterior probability > 0.5) and one PD patient was misclassified as control (posterior probability < 0.5). This yielded a final sensitivity of 95%, a specificity of 70% and an accuracy of 86.7%.

3.3 Paper II: Classification of iRBD and Parkinson's disease patients based on eye movements during sleep

The previous study demonstrated that analysis of EMs during sleep and EMG activity measured at the EOG side both hold potential of being biomarkers of PD. This study focuses on EMs alone and aims at illustrating the timely distribution and course of EMs. There exist no exact criteria for duration, amplitude or frequency of an EM in the PSG, and no clear criteria for when an EM is rapid or slow, or when they can be stated as present or absent. Therefore no golden standard exist, and therefore no standard can be used to develop an EM detector. Instead, it was chosen to evaluate EMs by use of a data-driven topic model. A topic model is a statistical model revealing "topics" or "themes" describing latent structures behind the generation of a collection of text documents. By applying the topic model on data describing EMs during sleep for control subjects, each sleep epoch will be presented as a mixture of three different states thought to be related to slow EMs, rapid EMs or no EMs. In this way, the timely course and the transitions between EM states can be illustrated for each subject and it can be investigated how well EMs from iRBD and PD patients fall into the EM states found based on control subjects. A topic model approach was chosen as it is a data-driven model suitable for subdivision of epochs into underlying themes, that may be unnoticeable by human scorers. Additionally, as one sleep epoch can contain rapid EMs, slow EMs as well as periods with no EMs, an indication of each epoch's EM state membership was preferred. Allowing one epoch to be presented as a mixture of states rather that a single state, entails that transitions between states can be illustrated as continuously passages rather than abrupt changes. A topic model called Latent Dirichlet Allocation (LDA) was chosen as it embraces all these ambitions.

The overall methodology of this study is seen in figure 3.4.

3.3.1 Methods: Topic modeling of eye movements

Subjects and recordings

A total of 40 subjects were included in this study. The same 30 subjects as in the first pilot study and addition 10 control subjects used for developing the topic model. The demographic data can be seen in table 3.4.

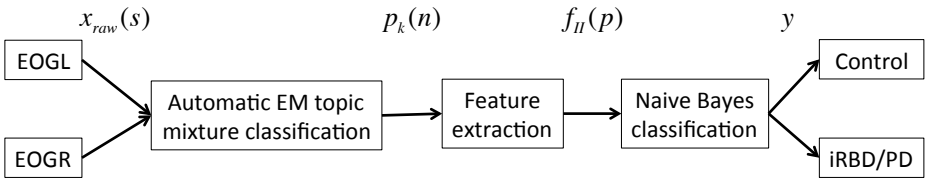


Figure 3.4: Overall methodology of the second study concerning eye movements (EMs) during sleep. Raw data $x_{raw}(s)$ from the left and right side EOG derivations were fed into an automatic EM topic mixture classification model, which outputs a topic mixture diagram indicating the probability mixtures $p_k(n)$ across K topics for each epoch n. For each topic mixture diagram, three features described in $f_{II}(p)$ were extracted and fed into a Naive Bayes classifier automatically classifying subjects as "control" or "patient".

	Controls (for train)	Controls (for test)	iRBD (for test)	PD (for test)
Total number	10	10	10	10
Male / Female	5 / 5	5 / 5	8 / 2	6 / 4
Age [years]	57.2 $\pm$ 8.1	59.8 $\pm$ 8.4	59.0 $\pm$ 14.2	63.2 $\pm$ 8.4

Table 3.4: Demographic data for the four groups analyzed in study II.

As in study I, the quality of the PSG recordings were individually evaluated and recordings with discontinued or artifact-contaminated channels were excluded. In this study, only the two EOG signals were used and they were both sampled with a sampling frequency of 256 Hz.

Generation of a topic model: Latent Dirichlet Allocation

Originally, topic models were used to investigate topics in a collection of text documents [5]. It is a statistical model that reveals underlying themes or topics describing the latent structures in a collection of text documents. Linking particular word combinations to the topic of a document, a topic model can identify the topics and give a topic mixture for each document based on statistics of the appearances of these words [5]. In this study, a common topic model approach called the Latent Dirichlet Allocation (LDA) was used. This approach assumes that a given document contains a combination of multiple topics, and it allows identification of concurrent topics, which are related to a small subset of words [5]. Specifically, LDA considers that each datapoint or word may belong to more than a single cluster or topic [3]. Considering that a collection of documents, or a corpus, consists of M documents all containing words drawn from the same overall fixed dictionary of V words indexed by $\{1, \ldots, V\}$, a document $\mathbf{w}$ can be described as a vector of N word indices,

$$\mathbf{w} = (w_1, \ldots, w_N), \quad w_n \in \{1, \ldots, V\}, \tag{3.6}$$

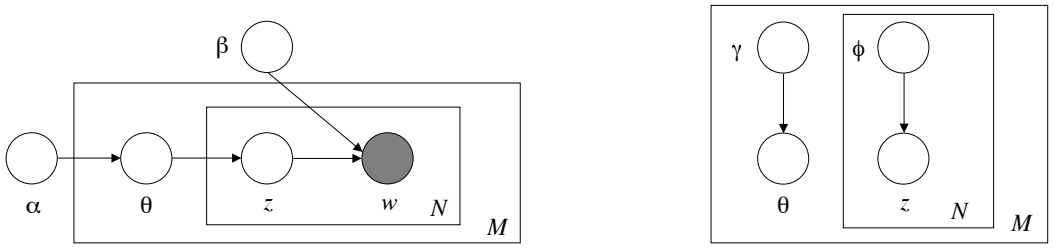
where N is the number of words in the the document. LDA assumes that each document may contain more than one topic, and therefore a distribution of topics $\boldsymbol{\theta}$ with $\sum_{k=1}^{K} \theta_k = 1$ exist for each document giving a description of the document in terms of its topic membership. To control for complexity, i.e. limit the number of topics to ensure that words are grouped and not considered individual topics, the Dirichlet prior is used to limit the number of topics active in each document [3]:

$$p(\boldsymbol{\theta} \mid \boldsymbol{\alpha}) = \text{Dirichlet}(\boldsymbol{\theta} \mid \boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{k=1}^{K} \alpha_k\right)}{\prod_{k=1}^{K} \Gamma(\alpha_k)} \theta_1^{\alpha_1 - 1} \dots \theta_K^{\alpha_K - 1}, \tag{3.7}$$

where $\boldsymbol{\alpha}$ is a vector of length K with components $\alpha_k > 0$, and $\Gamma(x)$ is the Gamma function.

In LDA, the generative process for each document $\mathbf{w}$ with N words is given by:

1. Choose a distribution of topics $\boldsymbol{\theta} \sim \text{Dirichlet}(\boldsymbol{\theta} \mid \boldsymbol{\alpha})$
2. For each word position w_n, $n = 1, \dots, N$:
   - Choose a topic z_n from the topic distribution: $z_n \sim p(z_n \mid \boldsymbol{\theta})$
   - Choose a word w_n from the word distribution of that topic: $w_n \sim p(w_n \mid z_n, \boldsymbol{\beta})$



(a) (b)

Figure 3.5: (a): Graphical model representation of LDA. The open circles are hidden variables and the filled one is known. The outer box (M) represents the repeated choice of documents and the inner box (N) represents the repeated choice of topics and words within a document. (b): Graphical model of the variational distribution used to approximate the posterior topic probabilities in LDA. Both figures are taken from [5].

A graphical model representation of LDA is given in figure 3.5(a). The open nodes are hidden variable and the filled node is a known variable. The word distribution $\boldsymbol{\beta}$ for each topic, and the parameters of the topic distribution $\boldsymbol{\alpha}$ are the parameters describing the model. These are corpus parameters and are sampled once per corpus. The document variables $\boldsymbol{\theta}$ are sampled once per document, whereas z and w are word-related variables sampled for each word in each document [5].

Training the LDA model corresponds to learning the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$. This is typically done through approximate inference algorithms, as the posterior distribution of the hidden

variables given by Bayes' rule,

$$p(\theta, \mathbf{z} \mid \mathbf{w}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{p(\theta, \mathbf{z}, \mathbf{w} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})}{p(\mathbf{w} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})} \quad \text{where} \tag{3.8}$$

$$p(\mathbf{w} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)} \int \left(\prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \right) \left(\prod_{n=1}^{N} \sum_{k=1}^{K} \prod_{j=1}^{V} \left(\theta_k \boldsymbol{\beta}_{kj} \right)^{w_n^k} \right) d\theta \tag{3.9}$$

is computationally intractable due to the coupling between $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ [5]. One example of an approximate inference algorithm is variational inference, which uses free parameters, γ and ϕ to find the tightest lower bound on the log-likelihood function. The problematic coupling between $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ arises due to the edges between θ, $\mathbf{z}$ and $\mathbf{w}$ seen in figure 3.5(a). Variational inference overcomes this by dropping the $\mathbf{w}$ nodes and the edges between θ and $\mathbf{z}$ to obtain a simplified graphical model with free variational parameters as seen in figure 3.5(b) [5]. Together γ and ϕ are a family of lower bounds giving the family of distributions on the latent variables θ and $\mathbf{z}$ characterized by

$$q(\theta, \mathbf{z} \mid \gamma, \phi) = q(\theta \mid \gamma) \prod_{n=1}^{N} q(z_n \mid \phi_n), \tag{3.10}$$

where γ is the Dirichlet parameter and $\phi_1, \ldots, \phi_N$ are the multinominal parameters. The optimization problem determining the value of the free variational parameters is now given by [5]

$$(\gamma^*, \phi^*) = \arg\min_{(\gamma, \phi)} D\left(q(\theta, \mathbf{z} \mid \gamma, \phi) \parallel p(\theta, \mathbf{z} \mid \mathbf{w}, \boldsymbol{\alpha}, \boldsymbol{\beta})\right), \tag{3.11}$$

which can be solved by use of an iterative fixed-point method and two update equations given by,

$$\phi_{nk} \propto \boldsymbol{\beta}_{kw_n} \exp\left\{\mathrm{E}_q\left[\log(\theta_k) \mid \gamma\right]\right\} \quad \text{and} \tag{3.12}$$

$$\gamma_k = \alpha_k + \sum_{n=1}^{N} \phi_{nk}. \tag{3.13}$$

The expectation in the multinomial update is given by:

$$\mathrm{E}_q\left[\log(\theta_k) \mid \gamma\right] = \Psi(\gamma_k) - \Psi\left(\sum_{j=1}^{K} \gamma_j\right) \tag{3.14}$$

where Ψ is the first derivative of the $\log\Gamma$ function which is computable via Taylor approximation [5]. As the optimization problem in equation 3.11 is conducted for fixed $\mathbf{w}$, γ^* and ϕ^* are dependent on $\mathbf{w}$ and could be written as $\gamma^*(\mathbf{w})$ and $\phi^*(\mathbf{w})$ instead. Given these values, the variational distribution $q(\theta, \mathbf{z} \mid \gamma^*(\mathbf{w}), \phi^*(\mathbf{w}))$ can thus be viewed as an approximation to the posterior distribution $p(\theta, \mathbf{z} \mid \mathbf{w}, \boldsymbol{\alpha}, \boldsymbol{\beta})$, or i.e. the joint distribution of the topic mixture θ and a set of K topics $\mathbf{z}$ given a set of N words $\mathbf{w}$, the Dirichlet parameters $\boldsymbol{\alpha}$ and the word distribution $\boldsymbol{\beta}$ for each of the topics. The optimization problem is solved for each document, and when applying a trained LDA topic model to new data, $\gamma^*(\mathbf{w})$ and $\phi^*(\mathbf{w})$ are approximated by use of equation 3.11 and the values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ learned trough training. Finally, the topic mixtures for a new document are approximated through iterations using the update equations stated in 3.12 and 3.13.

Generating topic model for EM states

Relating topic modeling to sleep data, a sleep epoch can be seen as a document and the words used for revealing the topics can be explained as variables or structures indicative for the different sleep states.

As this study focuses on EMs, the number of topics was set to $K = 3$ to reflect the three states (slow, rapid and no EMs), that are believed to be the overall states EMs can be in during sleep. In order to train a LDA topic model, a transformation from EOG signals to "word" is needed. As described in last section, each document, or sleep epoch, has to be described by a vector $\mathbf{w}$ In figure 3.6 is given an overview of the transformation of raw EOG data to a word vector $\mathbf{w}$. The words used for describing the EM topics were computed from the two EOG signals. Initially, both signals were bandpass filtered to focus on the frequencies in the range 0.3-10 Hz. The signals were divided into non-overlapping segments of length L s, for which three word features were computed:

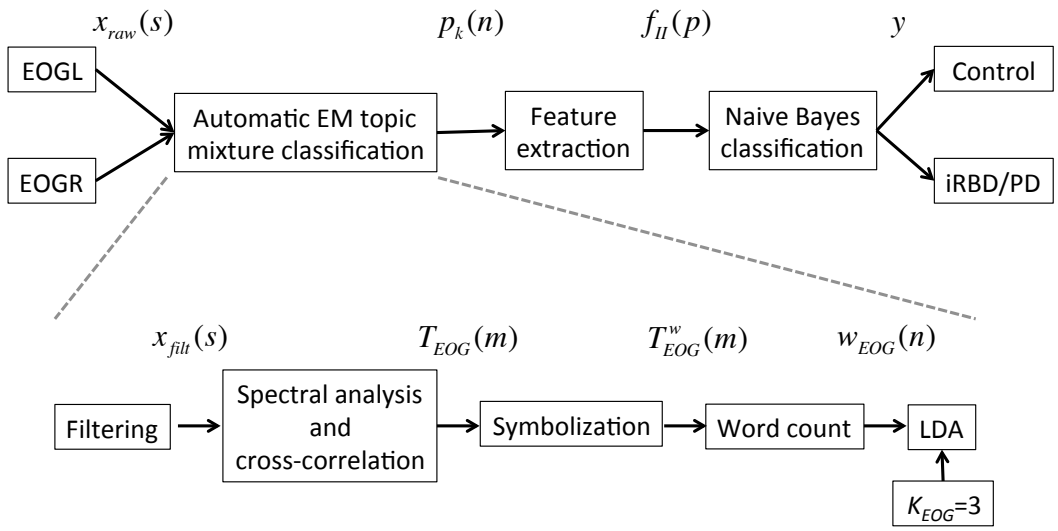
$$T_{EOG}(m) = \begin{bmatrix} X_{ll}(m) \\ X_{rr}(m) \\ X_{lr}(m) \end{bmatrix} \quad \text{where} \quad X_{lr}(m) = \sqrt{\frac{\sigma_{lr}(m)}{\sigma_{ll}^2(m)\sigma_{rr}^2(m)}} - \overline{X_{lr}} \tag{3.15}$$

where m denoted the segment index and X_{ll} and X_{rr} represents the spectral power computed by the fast Fourier Transform (FFT) below 5 Hz in the left and right EOG signal segment, respectively. The variance of the left and right EOG signal segment is denoted σ_{ll}^2 and σ_{rr}^2, respectively, and σ_{lr} denotes the covariance of the two segments. The normalized cross-correlation coefficient X_{lr} will obtain negative values when EMs are present (EOG signals anti-correlated), positive values when EEG artifacts such as delta waves are present (EOG signals correlated) and values close to zero when background EOG activity dominates the segment (EOG signals uncorrelated). To normalize X_{lr} across recordings, the values were aligned around zero by subtracting the subject-specific median $\overline{X_{lr}}$.

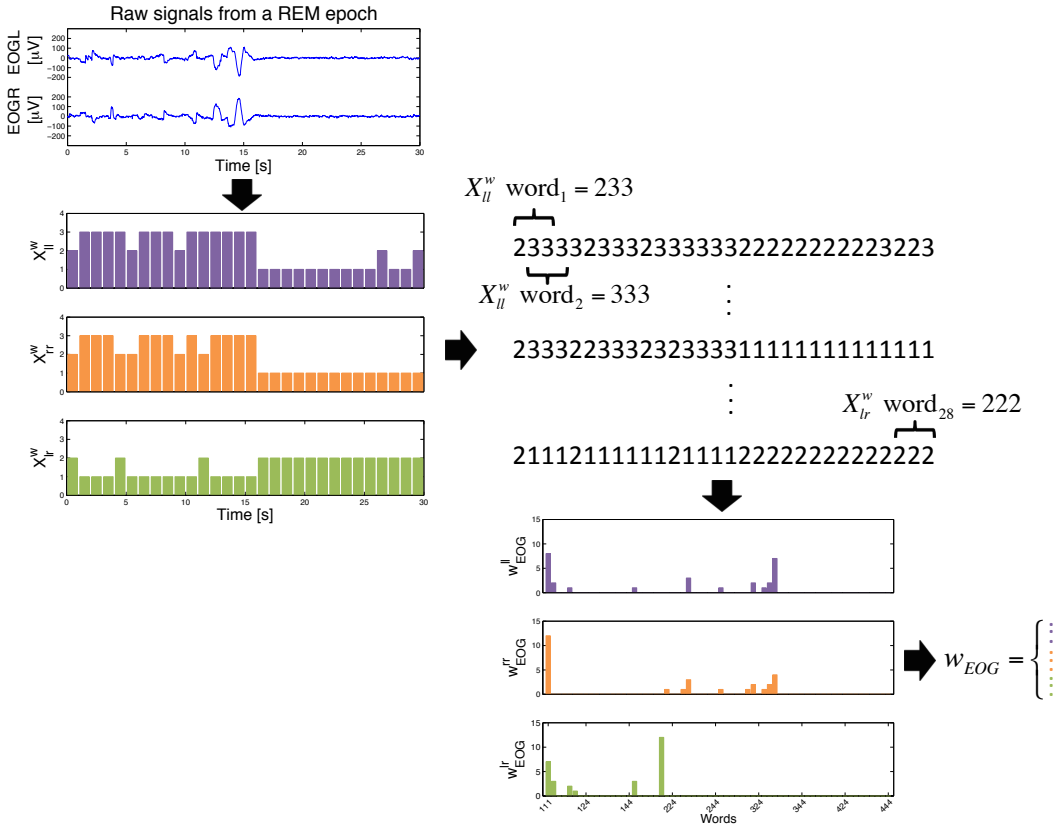
The word feature vector, $T_{EOG}(m)$ was converted into words by discretizing the values on a per-subject basis. The word features X_{ll} and X_{rr} were given the symbols 1 to 4 based on boundaries set at each quartile $q_{0.25}$, $q_{0.50}$, $q_{0.75}$ for the full range of feature values for that specific subject, described as,

$$X_{ll}^w = \begin{cases} 1 \text{ if } X_{ll} \leq q_{0.25}^l \\ 2 \text{ if } q_{0.25}^l < X_{ll} \leq q_{0.50}^l \\ 3 \text{ if } q_{0.50}^l < X_{ll} \leq q_{0.75}^l \\ 4 \text{ if } q_{0.75}^l < X_{ll} \end{cases} \quad \text{and} \quad X_{rr}^w = \begin{cases} 1 \text{ if } X_{rr} \leq q_{0.25}^r \\ 2 \text{ if } q_{0.25}^r < X_{rr} \leq q_{0.50}^r \\ 3 \text{ if } q_{0.50}^r < X_{rr} \leq q_{0.75}^r \\ 4 \text{ if } q_{0.75}^r < X_{rr} \end{cases} \tag{3.16}$$

where q^l and q^r denote quartiles based on values from X_{ll} and X_{rr}, respectively. The cross-correlation word feature X_{lr} was given symbols 1 to 4 based on boundaries set at $[-0.7, 0, 0.7]$



(a)



(b)

Figure 3.6: (a): Schematically presentation of the conversion of the raw EOG signals, $x_{raw}(s)$, to vectors, $w_{EOG}(n)$, describing the words of each sleep epoch n. The vectors are given as input to the Latent Dirichlet Allocation (LDA) model, which outputs the topic probability mixtures $p_k(n)$ across the $K_{EOG} = 3$ topics. (b): Graphic illustration of the conversion of raw EOG signals in a REM epoch to a vector $\mathbf{w}_{EOG}$ describing the words of that epoch. In this illustration the power values X_{rr} and X_{ll} and the cross-correlation values X_{lr} were given the symbols 1 to 4 and a word length was set to be $W = 3$.

for all subjects, described as,

$$X_{lr}^{w} = \begin{cases} 1 \text{ if } X_{lr} \leq -0.7 \\ 2 \text{ if } -0.7 < X_{lr} \leq 0 \\ 3 \text{ if } 0 < X_{lr} \leq 0.7 \\ 4 \text{ if } 0.7 < X_{lr} \end{cases} . \tag{3.17}$$

The boundary values were set based on trial-and-error of best catching the EMs (values below -0.7) and the EEG artifacts measured at the EOG sites (values above 0.7). X_{ll}^{w}, X_{rr}^{w} and X_{lr}^{w} together described the symbolized word feature vector T_{EOG}^{w}, as

$$T_{EOG}^{w}(m) = \begin{bmatrix} X_{ll}^{w}(m) \\ X_{rr}^{w}(m) \\ X_{lr}^{w}(m) \end{bmatrix}. \tag{3.18}$$

Different word lengths were tried ($W = 2, 3, 5$) resulting in EOG words to be presented by either one of all combinations of W succeeding symbols of 1 to 4. A sleep epoch of 30 seconds with a number of $\frac{30}{L}$ segments yields a total of 4^W available words for each word feature thereby giving a total of 3×4^W different words. For each sleep epoch, counts of the available words were performed, yielding a specific fingerprint of each epoch, here described by a vector $\mathbf{w}_{EOG}(n)$ where W=3:

$$\mathbf{w}_{EOG}(n) = \begin{bmatrix} \mathbf{w}_{EOG}^{ll}(n) \\ \mathbf{w}_{EOG}^{rr}(n) \\ \mathbf{w}_{EOG}^{lr}(n) \end{bmatrix} = \begin{bmatrix} \#\text{ ”111” in epoch } n \text{ of } X_{ll}^{w} \\ \#\text{ ”112” in epoch } n \text{ of } X_{ll}^{w} \\ \vdots \\ \#\text{ ”444” in epoch } n \text{ of } X_{ll}^{w} \\ \#\text{ ”111” in epoch } n \text{ of } X_{rr}^{w} \\ \#\text{ ”112” in epoch } n \text{ of } X_{rr}^{w} \\ \vdots \\ \#\text{ ”444” in epoch } n \text{ of } X_{rr}^{w} \\ \#\text{ ”111” in epoch } n \text{ of } X_{lr}^{w} \\ \#\text{ ”112” in epoch } n \text{ of } X_{lr}^{w} \\ \vdots \\ \#\text{ ”444” in epoch } n \text{ of } X_{lr}^{w} \end{bmatrix}, \tag{3.19}$$

where $\#$ denotes the total number. The topics are assumed to be defined by certain distribution of the available words ($\boldsymbol{\beta}$ in previous section), and using the word counts for each sleep epoch, a distribution over topics is derived ($\boldsymbol{\theta}$ in previous section). Specifically, LDA outputs the posterior probability $p_k(n)$ for EM topic k in epoch n.

Training a LDA topic model lies in learning the distribution of words $\boldsymbol{\beta}$ for each topic, and the parameter $\boldsymbol{\alpha}$ of the topic distribution, and in this study all sleep data from ten control subjects was used to train a general topic model. The trained topic model was applied on the three test

groups (see table 3.4) yielding a topic mixture diagram holding the posterior probabilities for the three EM states in each sleep epoch from each of the test subjects.

Feature extraction

For each test subject, three features were extracted from the topic diagrams. The features are summarized in table 3.5 and are for each subject p described by,

$$f_{II}(p) = \begin{bmatrix} f_{certainty}(p) \\ f_{fragmentation}(p) \\ f_{stability}(p) \end{bmatrix}. \tag{3.20}$$

Feature name	Feature explanation	Total no per subject EOG
Certainty	Normalized number of epochs with any topic probability higher than t	1
Fragmentation	Normalized number of transitions from one topic to another topic	1
Global stability	Normalized number of epochs in a period of epochs where the topic is the same	1

Table 3.5: Overview of the feature groups computed in study II. The threshold t denotes when an epoch was counted as certain. Different values were tried ($t = [0, 0.025, \ldots, 1]$) and the best was chosen as the one giving the highest Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) when classifying the 30 test subjects.

The certainty feature $f_{certainty}$ was designed to capture the trait that control subjects show more epochs with high topic probability compared to the patients, i.e. that they show a more "clear" sleep outlook as they have distinct differences between the EM states. The fragmentation feature $f_{fragmentation}$ was designed to capture the trait that patients show a more abrupt and fragmented pattern of EMs compared to controls. Finally, the stability feature $f_{stability}$ was designed to capture the trait that patients show a decreased ability to sustain a given EM state and thereby stay less time in any state before switching to another state compared to control subjects. The three features are described as,

$$f_{certainty} = \frac{1}{N}\sum_{n=1}^{N} 1_{\{p_k(n)>t\}} \tag{3.21}$$

$$f_{fragmentation} = \frac{1}{N}\sum_{n=1}^{N-1} 1_{\left\{\arg\max_k p_k(n) \neq \arg\max_k p_k(n+1)\right\}} \tag{3.22}$$

$$f_{stability} = \frac{1}{S}\sum_{s=1}^{S} L_s \quad \text{with} \quad L_s = \frac{l_s - \min(l_s)}{\max(l_s) - \min(l_s)} \tag{3.23}$$

where N is the subject-specific total number of epochs, $p_k(n)$ defined the probability of EM topic k in epoch n and 1 is an indicator function, and t defined when an epoch is counted

as certain. The best value for $t = [0, 0.025, ..., 1]$ was chosen as the one giving the highest Area Under The Receiver Operating Characteristic (ROC) Curve (AUC) when classifying the 30 test subjects using the leave-one-subject out approach. In $f_{stability}$, s is a period in which the dominant topic is maintained and the vector l_s expresses the number of epochs in such periods.

As the topic diagrams depend on the initialization of the LDA model, it was noticed that the features slightly changed across runs, and the final features used for classification was computed as the mean across 20 different runs on the test data.

Classification

Using the leave-one-subject-out classification approach, a standard Naive Bayes (NB) classifier was used to classify the test subjects as "control" or "patient". A NB classifier is a simple probabilistic classifier, which assigns an observation x to the class C_g with the highest posterior probability, i.e. it assigns the label y based on,

$$y = \underset{g \in \{1,...,G\}}{\arg\max}\, p\left(C_g\right) \prod_{i=1}^{n} p\left(x_i \mid C_g\right), \tag{3.24}$$

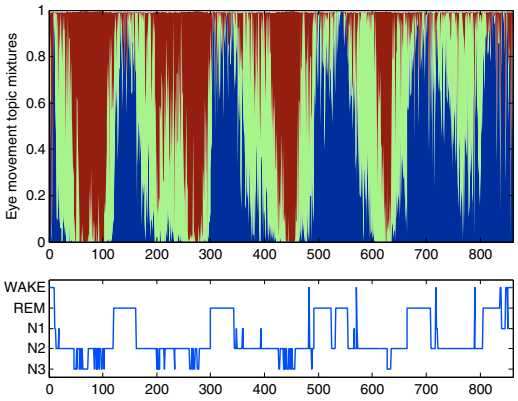
where g is the class index, $p\left(C_g\right)$ is the prior probability of class g and $p\left(x \mid C_g\right)$ is the likelihood function. By doing so, a NB classifier assigns x to the class with the minimum expected loss.

Different classification schemes were tried: The three features were given as input one by one, combined in pairs or all together, and the NB classifier with the highest AUC was chosen as the final one. The NB classifier was chosen for simplicity, and to make a brief evaluation of the features' capability to separate patients and controls.

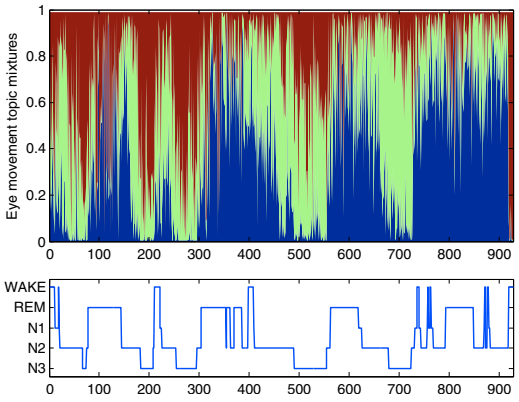
3.3.2 Results

The final topic model performing the best according to AUC after NB classification was found to include a word length of $W = 3$ and a segment length of $L = 1$ s. In figure 3.7 is seen examples of topic mixture diagrams from a control subject, an iRBD patient and a PD patient. Each vertical colored bin represents a sleep epoch, and the amount of each color represents the individual topic probability. The characteristics of the sleep evolution is captured although the model only includes information about EMs. The blue topic could be interpreted as having something to do with rapid EMs in REM sleep, the green topic could be linked to slow EMs and the red topic could be linked to states where no EMs are present.

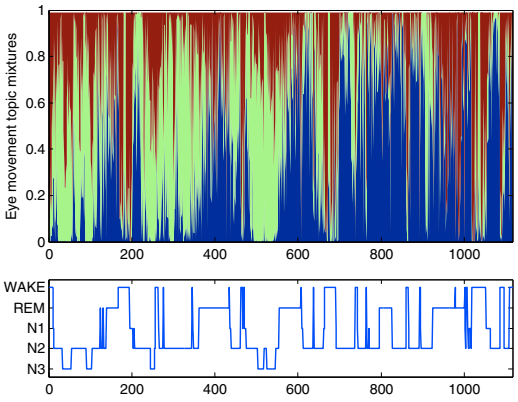
In figure 3.8 is the decision boundary obtained from the best NB classification illustrated by the colors gray (classified as "patient") and white (classified as "control"). The best model includes $f_{certainty}$ with a threshold of $t = 0.75$ and $f_{stability}$ and it obtained a sensitivity of 95%, a specificity of 80% and an accuracy of 90%.



(a) A control subject.



(b) An iRBD patient.



(c) A PD patient.

Figure 3.7: Examples of topic mixture diagrams expressing mixtures of eye movement (EM) topics. Each vertical bin represents a sleep epoch of 30 s and the amount of color in each bin represents the individual topic probability. For comparison, the manually scored hypnograms are provided below each diagram.

3.4 Conclusive remarks

Study I:

In the first study, EOG was found to hold indicative information of NDD. By use of the wavelet decomposition technique, features reflecting energy in different frequency bands were extracted for each sleep epoch. The means and standard deviations across all sleep epochs were computed to make a brief assessment of whether patients with iRBD or PD had values deviating from controls. Four features were found to be the most indicative ones, whereof two reflected energy in the frequency range (1-2 Hz) considered to be related to EMs during sleep. Specifically, these two features reflected the mean and the standard deviation of the percentage energy in the frequency range 1-2 Hz across all sleep epochs, thereby only reflecting the overall differences in EMs between controls and iRBD/PD patients. The study emphasizes that artifacts such as baseline drift as well as EEG activity measured at the EOG side might have had impact on the result.

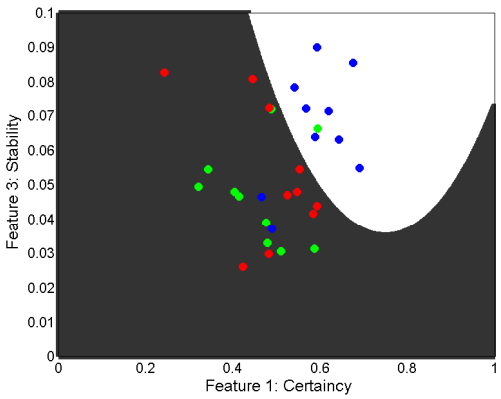


Figure 3.8: The best Naive Bayes classification result was based on $f_{certainty}$ and $f_{stability}$. A blue dot illustrates a control subject, a green dot illustrates an iRBD patients and a red dot illustrates a PD patient. The white and gray areas illustrate where the subjects are classified as controls and patients, respectively. One iRBD patients and two controls were misclassified yielding a specificity of 95%, a sensitivity of 80% and an accuracy of 90%.

Study II:

In the second study, the analysis was focused on EMs alone, and a data-driven topic model was developed independently of manually sleep scorings. The topic model was found to be helpful when evaluating EMs during sleep, and extracted features reflecting the overall patterns of EMs during sleep were found to be indicative of patients with iRBD or PD.

Overall:

Based on the two studies, it can be concluded that patients with iRBD and patients with PD reflect abnormal form and/or timely distribution of EMs during sleep. More focused analysis is needed to further investigate how the neurodegeneration effect the EM morphology (the rapidness, the amplitude etc.) as well as the EM controlling mechanisms, i.e. when are rapid and slow EMs present and are EMs limited to certain periods as in controls or are they distribution randomly across the night.

Due to neuroanatomy, EMs during sleep are very likely to possess indicative information about neurodegeneration, but as no standard exist for manually scoring start and stop for EMs, it is difficult to confirm their presence or absence. Consequently, it is hard to certify that the described EOG analyses only reflect EMs. Although limited, artifacts such as baseline drift, EEG delta waves, EEG micro-sleep structures and EMG activity recorded at EOG sites could all have non-neglectable influence on the results found.

Overall, this research area conclude:

- that EMs during sleep hold potential as a PD biomarker.
- that a data-driven topic model describing EMs during sleep can be helpful when analyzing and automatically classifying controls and patients with iRBD or PD.

4

Sleep stage switching and stability

Objective *Sleep disturbances are frequent complains in patients with NDDs, including PD. The control of sleep and wakefulness relies on multiple interacting switching mechanisms mainly regulated by neurons located in the brainstem and midbrain areas [7]; all of which are at risk of being affected in early stages of PD [6]. A malfunction or destruction of any of the switching mechanisms might be observed in the sleep architecture as either unstructured transitions, and/or abnormal amount of time spent in the individual stages. Sleep from patients with PD or iRBD is altered due to the neurodegeneration making it difficult to follow the standard for sleep scoring. This chapter is composed upon Paper III and IV, and serves to show the advantages of using topic modeling when analyzing sleep from patients with NDDs.*

4.1 Background

Shifts between sleep stages and transitions from sleep into wakefulness are controlled by switching mechanisms regulated by several neurons in the brainstem and midbrain areas as further described in section 2.2.3. The model suggests sleep-wake and NREM-REM flip-flop switches to be mutually dependent making them capable of generating complete and rapid transitions between wake and sleep stages.

Despite the mutually inhibitory loops involved in the two switching mechanisms, if either side of the two loops is weakened or injured, unwanted instability can occur in either of the states, irrespective of which side is damaged [74]. Neurodegeneration affecting the controlling area might therefore be observed in the sleep architecture as either unstructured transitions and/or abnormal amount of time spent in the individual stages.

Manually scoring of sleep in iRBD and PD patients is challenged with altered sleep characteristics such as fewer or abnormal sleep spindles (SS) [15], changes in slow-wave characteristics [47], frequency slowing [64], abnormal eye movements (EMs) [14] [16] and presence of REM sleep without atonia (RSWA) [41] [62]. The inter-rater reliability of scoring sleep in this pathology has been indicated to be low [21] [39], mainly because the scoring standard simply does not fit to altered sleep where stage-dependent characteristics can not be recognized. Sleep analysis would therefore benefit from a data-driven method that could look deeper into

the data and recognize the important characteristics in a robust and consistent way. In that way, data would lead the analysis independently of manual scorings.

The EMG activity is enhanced in iRBD and PD patients, suggesting that the neurodegenerative process occurs in the region of the brainstem generating REM atonia [7]. In other words, there is strong evidence that the neurodegeneration has affected the descending branch of neurons inhibiting the skeleton muscles during REM sleep. To avoid influence from the EMG abnormalities, the data-driven models developed are based on EOG and EEG features alone, hereby focusing on the characteristics affected by the ascending cortical parts of the sleep-wake and the REM-NREM sleep switches.

4.1.1 Research hypothesis

The main hypothesis for this research area was that patients with iRBD and PD express altered transitions between wake and sleep stages compared to controls. Alterations in several micro-sleep structures challenge the manually scoring of sleep in this pathology, which makes the manual scoring questionable when investigating the sleep stage transitions. Traditional sleep scoring assigns one sleep stage per sleep epoch, which does not reflect the various micro-sleep events in the epochs. By allowing each sleep epoch to be described by a mixture of sleep stages, the timely course of sleep and the transitions between sleep stages can be described in much more detail. These intentions were achieved by addressing sleep analysis by a data-driven topic modeling approach.

The aim with this research area was to:

- develop usable data-driven models based on EEG and/or EOG data from control subjects.
- use the data-driven models to analyze how well sleep data from iRBD and PD patients fit into the standard topics identified from control subjects.
- use the data-driven models to extract features reflecting sleep stage transitions, sleep stage stability and amount of time spent in each sleep stage.
- investigate the robustness of features for indication of NDD by including patients with a motor disturbance but not an NDD; i.e. patients with periodic leg movement disorder (PLMD).

4.2 Paper III: Data-driven modeling of sleep EEG and EOG reveals characteristics indicative of pre-Parkinson's and Parkinson's disease

In this study a data-driven topic model was developed based on EEG and EOG data from control subjects. The study evaluates the diagnostic value of features reflecting sleep characteristics such as the stability, fragmentation and distribution of sleep stages in patients with iRBD or PD. By addressing sleep analysis with a topic modeling approach, this study does not attempt to match the manually scored sleep stages, but aims at identifying topics in the EEG and EOG and thereby capture latent diversities between subjects with and without neurodegeneration.

The overall methodology of this study is seen in figure 4.1

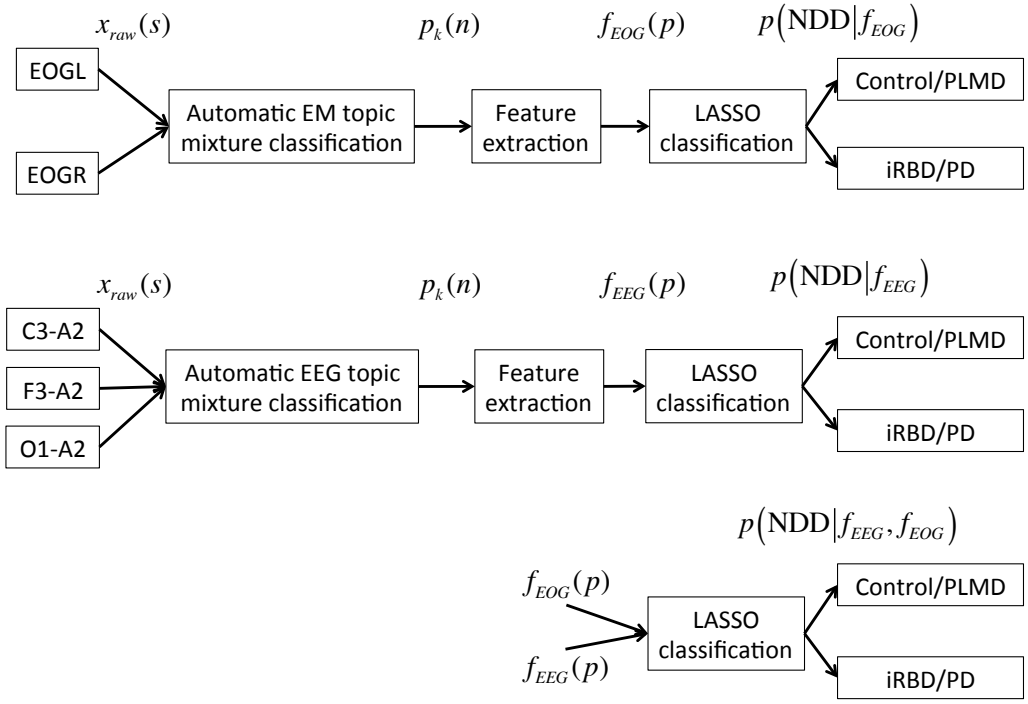


Figure 4.1: Overall methodology of study III dealing with sleep stability, fragmentation and sleep stage distribution as determined by a data-driven topic model approach. EEG and EOG data was used separately to build two topic models; an EEG topic mixture model and an EM topic mixture model, respectively. Features describing sleep EEG and EOG characteristics, $f_{EEG}(p)$ and $f_{EOG}(p)$, were extracted from the obtained topic mixture diagrams $p_k(n)$. Three Lasso regularized logistic regression models were trained; one given EEG features, one given EOG features and one given both as input. Indicative features were selected and used to classify the subjects by evaluating the posterior probability $p(NDD \mid f)$ of belonging to the "NDD" class. The "NDD" class holds the iRBD and PD patients and the "non-NDD" class holds controls and patients with periodic leg movement disorder (PLMD).

4.2.1 Methods: Topic modeling of EOG and EEG data

Subjects and recordings

A total of 125 subjects were enrolled in this study. Of these, 36 were PD patients, 31 were iRBD patients, 25 were patients with PLMD and 33 were control subjects. All patient evaluations included a comprehensive medical and medication history and a PSG according to the AASM standard [34]. Patients treated with medication known to effect sleep stages (antidepressants, antipsychotics, hypnotics) were excluded, although dopaminergic treatments were continued. The PLMD patients were included as an additional control group to stress the model as these patients show motor abnormalities, but should not be confused with patients with NDD. The PLMD patients did not show any signs of neurodegeneration or RSWA. No PSG findings or NDD-related symptoms were reported for this group and they were considered as solely PLMD patients. The control subjects had no history of movement disorder, dream-enacting (DE) behavior or other previously diagnosed sleep disorders. The quality of the PSG data was individually evaluated, and recordings were excluded if the analyzed channels were disconnected or continuously contaminated with artifacts.

The data was split into three groups: one for developing the topic models (10 control subjects), another for training Lasso regularized logistic regression models for classifying NDD patients (16 subjects from each group) and a third for final validation of the classifications (7 control subjects, 9 PLMD, 15 iRBD and 20 PD patients). The datasets were designed to be matched by age with no further knowledge of the subjects. The building dataset was, however, designed to be balanced in age and sex. The demographic data of the datasets are seen in table 4.1, and the distribution of the manually scored sleep stages are summarized in table 4.2.

Patient group	Total number (M/F)	Age [years]	BMI [kg/m^2]	SE [%]	BT [min]	LM index	Disease duration [years]
Data used for building topic models							
Controls	10(5/5)	56.0±8.4	23.8±2.7	89.4±5.5	480±52.9	38.6±43.8	NA
Training dataset							
Controls	16(5/11)	56.7±10.2	23.1±2.6	87.7±8.3	490±72.3	21.6±13.7	NA
PLMD	16(9/7)	56.9±11.8	25.1±3.5	86.5±6.1	434±32.7	60.4±42.3	NA
iRBD	16(13/3)	62.9±8.6	25.8±3.3	87.4±5.7	485±104.2	44.4±32.0	NA
PD	16(11/5)	65.4±6.4	26.4±2.6	78.5±13.5	462±87.9	47.1±47.6	3.6±4.5
Validation dataset							
Controls	7(2/5)	56.9±7.4	23.1±2.1	82.7±15.5	470±103.6	10.8±8.5	NA
PLMD	9(4/5)	57.0±12.1	26.7±4.3	85.0±9.8	428±73.9	49.7±25.1	NA
iRBD	15(13/2)	63.4±5.8	25.4±3.1	80.8±8.6	486±94.0	32.8±26.4	NA
PD	20(13/7)	65.1±6.9	25.0±3.4	76.3±13.7	430±66.6	29.4±30.3	8.8±3.5

Table 4.1: Demographic characteristics of the four groups used to develop, train and validate the topic models. The characteristics are indicated by means and standard deviations ($\mu \pm \sigma$). SE: Sleep efficiency; BT: Time in Bed; LM: Leg movements.

Patient group	Wake [No. (%)]	REM [No. (%)]	N1 [No. (%)]	N2 [No. (%)]	N3 [No. (%)]	Sum [No. (%)]
Data used for building topic models (epochs were selected to have an approximately equal number in each stage)						
Controls	642 (20)	700 (22)	617 (19)	700 (22)	597 (18)	3256 (100)
Training dataset						
Controls	1887 (12)	3114 (20)	1292 (8)	7091 (45)	3201 (15)	15685 (100)
PLMD	1836 (13)	2767 (20)	1090 (8)	6010 (43)	2170 (16)	13873 (100)
iRBD	2012 (13)	2923 (19)	1637 (11)	6789 (44)	2149 (14)	15510 (100)
PD	3113 (21)	1652 (11)	1867 (13)	6251 (42)	1889 (13)	14772 (100)
Validation dataset						
Controls	1137 (17)	1136 (17)	683 (10)	2890 (44)	733 (11)	6579 (100)
PLMD	1141 (15)	1204 (16)	732 (10)	3674 (48)	951 (12)	7702 (100)
iRBD	2967 (20)	2098 (14)	1106 (8)	5880 (40)	2535 (17)	14586 (100)
PD	4057 (24)	2464 (14)	1521 (9)	6523 (38)	2625 (15)	17190 (100)

Table 4.2: The distribution of sleep stages as determined by the manually scored hypnogram used in the development, training and validation of the EEG and EOG topic models.

Development of an EEG and an EOG topic model

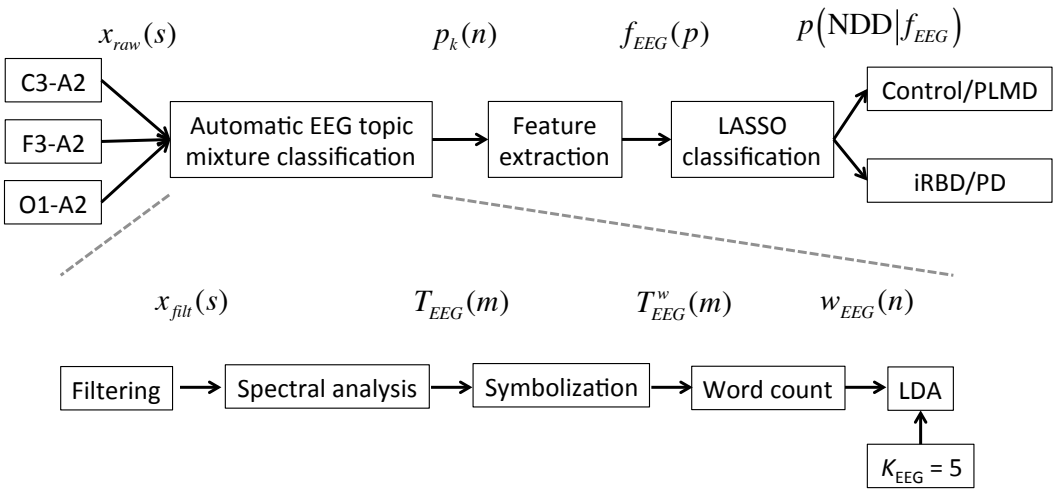
Sleep data from 10 control subjects was used to build two general topic models; one based solely on EOG and one based solely on EEG. As in paper II described in section 3.3, the LDA topic model was used to build the models. Again, conversion of EEG and EOG data into words was needed, thereby describing each sleep epoch as a word vector of EEG words, $\mathbf{w}_{EEG}(n)$, and a word vector of EOG words, $\mathbf{w}_{EOG}(n)$.

A schematic and a graphic illustration of the conversion of raw EEG data to the word vector $\mathbf{w}_{EEG}(n)$ fed into LDA is seen in figure 4.2.

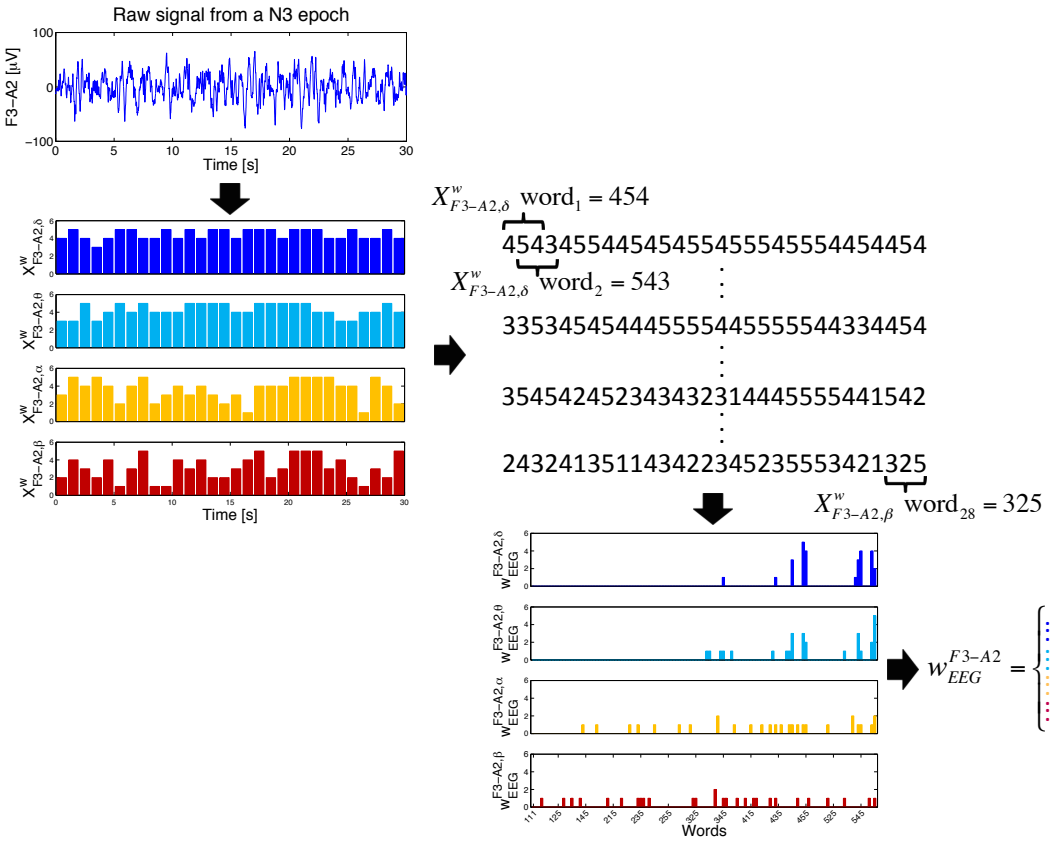
The EEG words were extracted from the C3-A2, F3-A2 and O1-A2 EEG derivations, all sampled with a sampling frequency of $f_s = 256$ Hz. After the signals were filtered forward and reverse with a bandpass filter with cutoff frequencies at 0.3 and 35 Hz, they were divided into non-overlapping 1-s segments. Four EEG word features were computed for each segment m, yielding a word feature vector $T_{EEG}(m)$ expressed as,

$$T_{EEG}(m) = \begin{bmatrix} X_{C3-A2}(m) \\ X_{F3-A2}(m) \\ X_{O1-A2}(m) \end{bmatrix} \quad \text{where} \quad X(m) = \begin{bmatrix} x_\delta(m) \\ x_\theta(m) \\ x_\alpha(m) \\ x_\beta(m) \end{bmatrix} \tag{4.1}$$

represents the spectral power in the clinical EEG frequency bands delta ($\delta : f <$ 4Hz), theta (θ : 4 Hz $\leq f <$ 8 Hz), alpha (α : 8 Hz $\leq f <$ 13 Hz) and beta (β : 13 Hz $\leq f <$ 30 Hz) computed by the fast Fourier transform (FFT) using zeropadding and a rectangular window function. The EEG word feature vector $T_{EEG}(m)$ was converted into EEG words by assigning the spectral power values symbols 1 to 5 based on subject-specific boundaries set at each quintile, $q_{0.20}$, $q_{0.40}$, $q_{0.60}$ and $q_{0.80}$, for the full range of spectral power values for that specific



(a)



(b)

Figure 4.2: (a): Schematically presentation of the conversion of raw EEG signals, $x_{raw}(s)$, to vectors $\mathbf{w}_{EEG}(n)$ describing the words of each sleep epoch n. The word vectors are given as input to the Latent Dirichlet Allocation (LDA) model, which outputs the probability mixtures $p_k(n)$ of the $K_{EEG} = 5$ topics. (b): Graphic illustration of the conversion of one epoch of raw EEG signal to a vector $\mathbf{w}^{F3-A2}_{EEG}$ describing the EEG words of the frontal EEG derivation. The same was done for the central and occipital derivation, and $\mathbf{w}^{C3-A2}_{EEG}$, $\mathbf{w}^{F3-A2}_{EEG}$ and $\mathbf{w}^{O1-A2}_{EEG}$ were combined into the final word vector $\mathbf{w}_{EEG}(n)$ as seen in figure 4.3.

subject, derivation and clinical band. The EEG word feature vector is described by,

$$T^w_{EEG} = \begin{bmatrix} X^w_{C3-A2} \\ X^w_{F3-A2} \\ X^w_{O1-A2} \end{bmatrix} \text{ where } X^w_d = \begin{bmatrix} x^w_\delta \\ x^w_\theta \\ x^w_\alpha \\ x^w_\beta \end{bmatrix} \text{ where } x^w_c = \begin{cases} 1 \text{ if } x_c \leq q^{d,c}_{0.20} \\ 2 \text{ if } q^{d,c}_{0.20} < x_c \leq q^{d,c}_{0.40} \\ 3 \text{ if } q^{d,c}_{0.40} < x_c \leq q^{d,c}_{0.60} \\ 4 \text{ if } q^{d,c}_{0.60} < x_c \leq q^{d,c}_{0.80} \\ 5 \text{ if } q^{d,c}_{0.80} < x_c \end{cases} \quad (4.2)$$

where $q^{d,c}$ denotes the quintile for the power values in the clinical band c in the EEG derivation d. The word length was set to be three, and an EEG word was thus defined as either one of all combinations of three succeeding symbols of 1-5. This was done for all three EEG derivations, and a final EEG fingerprint of a sleep epoch was thus presented as the word count across the $3 \times 4 \times 5^3 = 1500$ available words (3 EEG derivations, 4 clinical bands, 5^3 word combinations), described by a vector $\mathbf{w}_{EEG}(n)$,

$$\mathbf{w}_{EEG}(n) = \begin{bmatrix} \mathbf{w}^{C3-A2}_{EEG}(n) \\ \mathbf{w}^{F3-A2}_{EEG}(n) \\ \mathbf{w}^{O1-A2}_{EEG}(n) \end{bmatrix} = \begin{bmatrix} \#\text{"111" in epoch } n \text{ of } X^w_{C3-A2,\delta} \\ \#\text{"112" in epoch } n \text{ of } X^w_{C3-A2,\delta} \\ \vdots \\ \#\text{"555" in epoch } n \text{ of } X^w_{C3-A2,\delta} \\ \#\text{"111" in epoch } n \text{ of } X^w_{C3-A2,\theta} \\ \vdots \\ \#\text{"555" in epoch } n \text{ of } X^w_{C3-A2,\theta} \\ \vdots \\ \#\text{"111" in epoch } n \text{ of } X^w_{C3-A2,\beta} \\ \vdots \\ \#\text{"555" in epoch } n \text{ of } X^w_{C3-A2,\beta} \\ \#\text{"111" in epoch } n \text{ of } X^w_{F3-A2,\delta} \\ \vdots \\ \#\text{"555" in epoch } n \text{ of } X^w_{F3-A2,\beta} \\ \#\text{"111" in epoch } n \text{ of } X^w_{O1-A2,\delta} \\ \vdots \\ \#\text{"555" in epoch } n \text{ of } X^w_{O1-A2,\beta} \end{bmatrix}, \quad (4.3)$$

where # denotes the total number. As in study II described in section 3.3, the EOG words were extracted from the left and right side EOG derivations, also sampled with a sampling frequency $f_s = 256$ Hz. After the signals were filtered forward and reversed with a bandpass filter with cutoff frequencies at 0.3 and 10 Hz, they were divided into non-overlapping 1-s segments. Three EOG word features were computed for each segment m, yielding a vector

$T_{EOG}(m)$ expressed as,

$$T_{EOG}(m) = \begin{bmatrix} X_{ll}(m) \\ X_{rr}(m) \\ X_{lr}(m) \end{bmatrix} \quad \text{where} \quad X_{lr} = \sqrt{\frac{\sigma_{lr}(m)}{\sigma_{ll}^2(m)\sigma_{rr}^2(m)}} - \overline{X_{lr}} \tag{4.4}$$

represents the normalized cross-correlation coefficient between the left and right EOG signal segment m. It is based on the variance of the left (σ_{ll}^2) and right (σ_{rr}^2) signal segment and the covariance (σ_{lr}) of the left and right signal segment. X_{ll} and X_{rr} represent the spectral power below 5 Hz in the left and right EOG signal segment, respectively, computed by FFT using zeropadding and a rectangular window. $\overline{X_{lr}}$ describes the subject-specific median. The transformation into EOG words was done by assigning the spectral power values symbols 1 to 3 based on subject-specific boundaries set at each tertile, $q_{0.33}$ and $q_{0.66}$, for the full range of spectral power values for that specific subject and derivation. X_{lr} were assigned symbols 1 to 3 based on boundaries set at [-0.7; 0.7] for all subjects. The symbolized EOG word feature vector T_{EOG}^w is described as,

$$T_{EOG}^w = \begin{bmatrix} X_{ll}^w \\ X_{rr}^w \\ X_{lr}^w \end{bmatrix} \text{ where } X_d^w = \begin{cases} 1 \text{ if } X_d \leq q_{0.33}^d \\ 2 \text{ if } q_{0.33}^d < X_d \leq q_{0.66}^d \\ 3 \text{ if } q_{0.66}^d \leq X_d \end{cases} \text{ and}$$

$$X_{lr}^w = \begin{cases} 1 \text{ if } X_{lr} \leq -0.7 \\ 2 \text{ if } -0.7 < X_{lr} \leq 0.7 \\ 3 \text{ if } 0.7 \leq X_{lr} \end{cases}, \tag{4.5}$$

where q^d is the tertile for the d EOG derivation. As in paper II, the boundaries [-0.7; 0.7] were set on a trial-and-error basis for the best separation of the data into either EMs (values below -0.7), background EOG (values between -0.7 and 0.7) or EEG artifacts measured at the EOG site (values above 0.7). The word length was set to two, and an EOG word was thus defined as either one of all combinations of two succeeding values of 1 to 3. A final EOG fingerprint of a sleep epoch was thus presented as the word count across the $3 \times 3^2 = 27$ available words (2 EOG derivations plus 1 cross-correlation measure, 3^2 word combinations), described by the EOG word vector $\mathbf{w}_{EOG}$,

$$\mathbf{w}_{EOG}(n) = \begin{bmatrix} \mathbf{w}_{EOG}^{ll}(n) \\ \mathbf{w}_{EOG}^{rr}(n) \\ \mathbf{w}_{EOG}^{lr}(n) \end{bmatrix} = \begin{bmatrix} \#\text{"11" in epoch } n \text{ of } X_{ll}^w \\ \#\text{"12" in epoch } n \text{ of } X_{ll}^w \\ \vdots \\ \#\text{"33" in epoch } n \text{ of } X_{ll}^w \\ \#\text{"11" in epoch } n \text{ of } X_{rr}^w \\ \vdots \\ \#\text{"33" in epoch } n \text{ of } X_{rr}^w \\ \#\text{"11" in epoch } n \text{ of } X_{lr}^w \\ \vdots \\ \#\text{"33" in epoch } n \text{ of } X_{lr}^w \end{bmatrix}, \tag{4.6}$$

where # denotes the total number. Data from 10 control subjects was used to build a general EEG topic model and a general EOG topic model. Approximately 70 sleep epochs from each AASM sleep stage were taken out randomly in between lights off and lights on from each of the ten control subjects. Approximating an equal representation of the AASM sleep stages equalized the changes for each AASM stage to be described by a topic. Not all subjects had 70 epochs of all stages, and the final number of epochs used in the building process is shown in table 4.2. Figure 4.3 illustrates how the EEG and EOG word vector for each of the epochs in the building dataset was combined into an EEG and EOG word matrix (i.e. to form the EEG and EOG "document corpus"). These matrices were each fed into LDA to built the general EEG and EOG topic model, respectively.

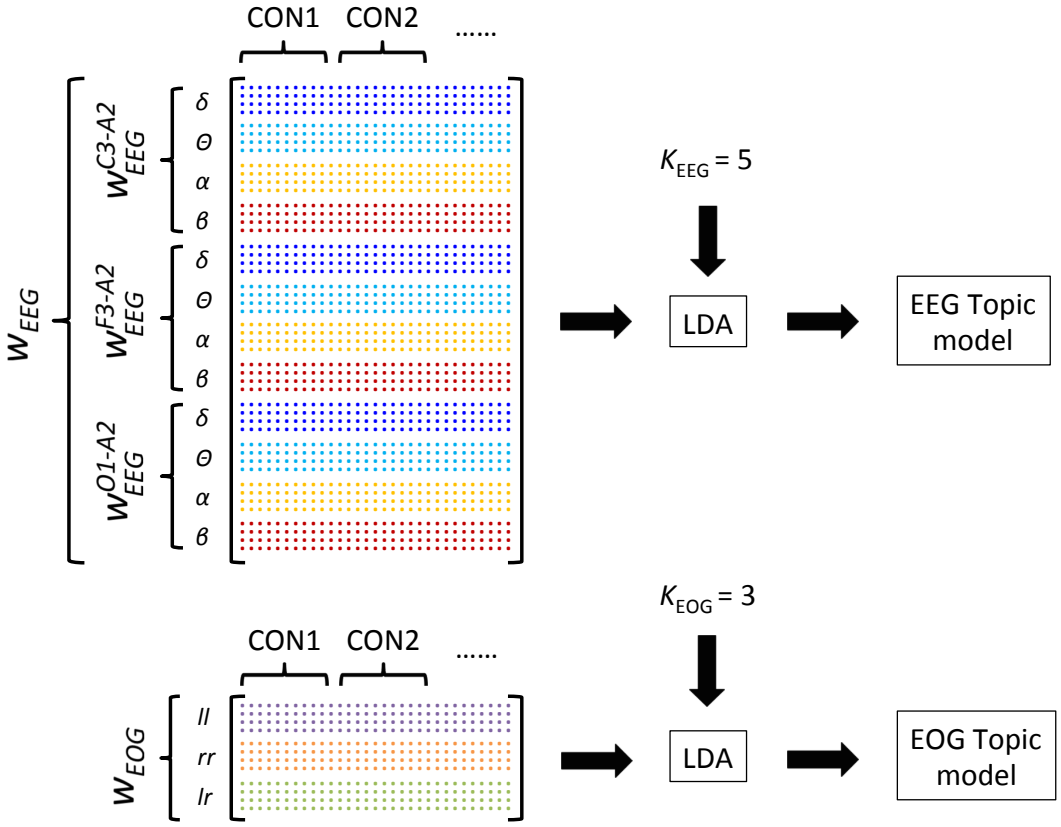


Figure 4.3: Illustration of how the EEG and EOG word vectors, $\mathbf{w}_{EEG}$ and $\mathbf{w}_{EOG}$, describing the EEG or EOG words for each sleep epoch, conform the input matrices, or in topic modeling language the "document corpora". Together with a fixed number of topics,K_{EEG} and K_{EOG}, the word matrices are given as input to the Latent Dirichlet Allocation (LDA) in order to built the general EEG and EOG topic model.

To reflect the number of AASM stages, the number of EEG topics was set to $K_{EEG} = 5$. To reflect the three major EMs during sleep (slow, rapid or no EMs), the number of EOG topics was set to $K_{EOG} = 3$.

For subjects included in the training or validation dataset, the general topic models were applied on all epochs between lights off and lights on, yielding an EEG topic diagram and

an EOG topic diagram for each subject describing the distribution over EEG or EOG topics in each sleep epoch, respectively. Specifically, LDA outputs the posterior probability $p_k(n)$ for sleep topic k in epoch n based on the distribution across words learned from the building dataset. In figure 4.4 is seen an example of an EEG and EOG topic mixture diagram from a control subject in the training dataset. Each vertical bin presents a sleep epoch and the amount of each of the colors in each bin represents the probability of each topic. EEG topics are illustrated by colors and EOG topics are presented by gray, white and black.

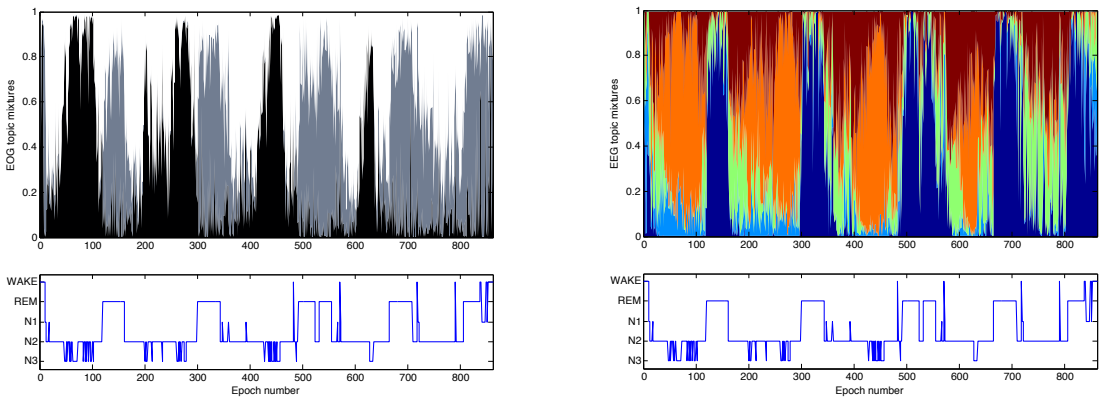


Figure 4.4: Examples of an EOG and an EEG topic mixture diagram from a control subject in the training dataset. A sleep epoch of 30 seconds is presented as a vertical colored bin, where each color represents a topic. The amount of color in each bin illustrates the probability of that topic. EEG topics are illustrated by colors and EOG topics are illustrated by gray, white and black. For comparison, the manually scored hypnogram is presented below the diagrams.

Feature extraction

Table 4.3 gives an overview of the features computed in this study, and the feature vector extracted from each of the two topic models can be described as,

$$f = \begin{bmatrix} f_{certainty} \\ f_{fragmentation} \\ f_{stability} \\ f^k_{T_{amount}} \\ f^k_{T_{stability}} \end{bmatrix}, \tag{4.7}$$

where k denotes the topic number. The certainty, fragmentation and global stability features were computed once for each topic model. The topic amount feature group $f^k_{T_{amount}}$ and the topic stability feature group $f^k_{T_{stability}}$ each comprise five features computed from the EEG topic diagrams and three features from the EOG topic diagrams, reflecting the number of topics for the diagrams. A total of nine features were thus extracted from the EOG topic diagrams and 13 features from the EEG topic diagrams.

Feature name	Feature explanation	Total no per subject EEG	EOG
Certainty	Normalized number of epochs with any topic probability higher than 0.90	1	1
Fragmentation	Normalized number of transitions from one topic to another topic	1	1
Global stability	Normalized number of epochs in a period of epochs where the topic is the same	1	1
Topic amount	Normalized number of epochs with a given topic, and with a topic probability higher than 0.70	5	3
Topic stability	Normalized number of epochs in a period of epochs with the same given topic	5	3

Table 4.3: Overview of the feature groups computed in study III.

The "certainty feature", "fragmentation feature" and the "global stability feature" were computed in a similar way as described in section 3.3.1 and were given by:

$$f_{certainty} = \frac{1}{N}\sum_{n=1}^{N} 1_{\{p_k(n)>t\}} \tag{4.8}$$

$$f_{fragmentation} = \frac{1}{N}\sum_{n=1}^{N-1} 1_{\left\{\arg\max_k p_k(n) \neq \arg\max_k p_k(n+1)\right\}} \tag{4.9}$$

$$f_{stability} = \frac{1}{S}\sum_{s=1}^{S} L_s \quad \text{with} \quad L_s = \frac{l_s - \min(l_s)}{\max(l_s) - \min(l_s)} \tag{4.10}$$

where N is the subject-specific total number of epochs, $p_k(n)$ is the probability of topic k for epoch n and s is a period of epochs for which the dominant topic is the same. In this study, $t = 0.9$ is the threshold value defining when an epoch is counted as certain and 1 is an indicator function. The threshold of $t = 0.9$ was decided as it was considered to be a reasonable high threshold for when an epoch should be counted as mainly concerning one topic. Also, by visual judgment of several threshold values in the range 0.6-1, this threshold seemed to be the best to separate controls and iRBD/PD patients. As in paper II, these features try to capture the characteristics that diagrams from non-NDD subjects show 1) more bins with a high probability (certainty feature), 2) more structure and less fragmentation (fragmentation feature) and 3) more subsequent epochs with the same dominant color (global stability feature) compared to the diagrams from the NDD patients. The non-NDD subjects encompasses the control subjects and the PLMD patients and the NDD patients hold the iRBD and PD patients.

The "topic amount feature" and "topic stability feature" were computed for each of the three EOG topics and five EEG topics and can be expressed as:

$$f^k_{T_{amount}} = \frac{1}{N}\sum_{n=1}^{N} 1_{\{p_k(n)>t_k\}} \tag{4.11}$$

$$f^k_{T_{stability}} = \frac{1}{N}\sum_{n=1}^{N-1} 1_{\left\{\arg\max_j(\, p_j(n)\,)=k \;\text{ and }\; \arg\max_j(\, p_j(n+1))=k\right\}} \tag{4.12}$$

where k is the topic number, $t_k = 0.7$ is a certainty threshold defining when an epoch is counted as a topic k epoch. Again, the threshold $t_k = 0.7$ was decided as it was considered to be the best to separate the groups based on visual judgment of several thresholds in the range 0.6-1. These features reflect the amount and stability of each of the three EOG topics and five EEG topics, and try to capture the traits that diagrams from non-NDD subjects 1) show changed amount of the different topics and 2) more often remain in a given topic for longer time compared to the diagrams from the NDD patients.

Feature ranking and classification

A subset of features was chosen using a logistic regression with Lasso regularization. This classification method was chosen as the coefficients are easy to interpret and the regularization forces the coefficients of irrelevant or redundant features to zero thereby yielding simple models less likely to overfit [33] [81]. In logistic regression a logistic sigmoid function is used to model the posterior probability of the *positive* class for the input variable $\mathbf{y}$:

$$p(\text{positive} \mid \mathbf{y}) = \frac{1}{1 + e^{-(\beta_0 + \boldsymbol{\beta}^T y)}} \tag{4.13}$$

where β_0 is the model offset and $\boldsymbol{\beta}$ is a vector of the variable coefficients. The inverse of the logistic function is the *logit* transformation expressed as,

$$\log \frac{p(\text{positive} \mid \mathbf{y})}{1 - p(\text{positive} \mid \mathbf{y})} = \beta_0 + \boldsymbol{\beta}^T \mathbf{y} \tag{4.14}$$

and the final classification is made by defining the decision boundary $\beta_0 + \boldsymbol{\beta}^T \mathbf{y} = 0$. The parameters in logistic regression models are often trained using maximum likelihood, maximizing the log-likelihood L expressed as:

$$L(\beta_0, \boldsymbol{\beta}) = \sum_{i=1}^{I} \left\{ z_i(\beta_0 + \boldsymbol{\beta}^T y_i) - \log\left(1 + e^{\beta_0 + \boldsymbol{\beta}^T y_i}\right) \right\} \tag{4.15}$$

where I is the number of observations and z_i is the classification output stating $z_i = 1$ for the *positive* class (NDD) and $z_i = 0$ for the *negative* class (non-NDD). The complexity of the model is controlled by introducing a regularization term $\lambda > 0$ yielding an optimization problem in the Lasso regularized logistic regression expressed as [81],

$$\arg\max_{(\beta_0, \boldsymbol{\beta})} \left\{ \sum_{i=1}^{I} \left[\mathbf{z}_i \left(\beta_0 + \boldsymbol{\beta}^T \mathbf{y}_i \right) - \log\left(1 + e^{\beta_0 + \boldsymbol{\beta}^T \mathbf{y}_i}\right) \right] - \lambda \sum_{d=1}^{D} \mid \boldsymbol{\beta}_d \mid \right\} \tag{4.16}$$

where D is the dimension of the input variable. Lasso regularization performs feature selection as some of the coefficients are driven to zero, thereby excluding the corresponding feature. As the regularization term λ increases, the number of nonzero components of $\boldsymbol{\beta}$ decreases and the model becomes more sparse [81]. Only features with nonzero components are included in the final model, and the size of the component value indicates each feature's impact in the final posterior probability output.

Sixteen subjects from each of the four groups were included in the training dataset, which was used to optimize the Lasso regularized logistic regression model using 8-fold cross validation. Three models were trained: (1) one in which all features were available initially, (2) one in which only the EOG features were available initially and (3) one in which only the EEG features were available initially. For each trained model, values for β and λ were obtained, thereby giving the optimal feature subset for each case. By training three models, the performance of the EOG and EEG features alone as well as in combination were evaluated.

4.2.2 Results

Figure 4.5 presents examples of obtained EEG and EOG topic diagrams from a control subject, a PLMD, an iRBD and a PD patient, respectively. The diagrams from the iRBD and PD patients were less structured, had a more fragmented profile, more abrupt transitions between topics and fewer epochs with a high certainty of either topic compared to the topic diagrams from control subjects or PLMD patients. The features were designed to detect these characteristics.

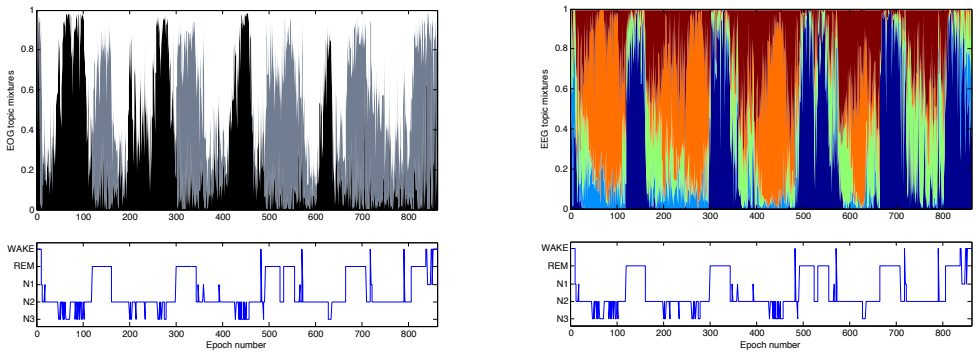
Visual comparison between the manually scored hypnograms and topic models show clear concordances, especially for the EEG topic diagrams. Table 4.4 gives an overview of the visual interpretation of the different EEG and EOG topics. Some of the topics were easy to link to one specific sleep stage and some were harder, as e.g. the EEG green and red topic and the EOG gray and white topic which were present in several sleep stages. Specifically, the black EOG topic may be related to the low-frequency high-amplitude EEG artifacts measured at the EOG during N3 sleep. The gray topic may be related to the rapid EMs during REM sleep and the white may include both the low-amplitude slow EMs and periods with no EMs.

Topic	EOG black	EOG gray	EOG white	EEG dark-blue	EEG light-blue	EEG green	EEG orange	EEG red
AASM stage	N3	REM/W	N1/N2	REM	W	N1/N2	N3	N1/N2

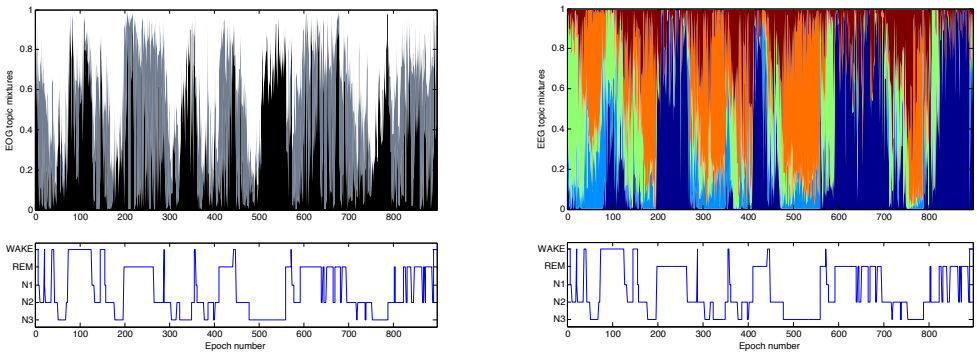
Table 4.4: Overview of the visual interpretation of the different EEG and EOG topics. Some topics were easy to link to one sleep stage and some topics were well presented during several sleep stages.

Figure 4.6 illustrates the feature values extracted from the EEG and EOG topic diagrams from each subject, and figure 4.7 shows the output from the final logistic regression models. The final models were expressed as,

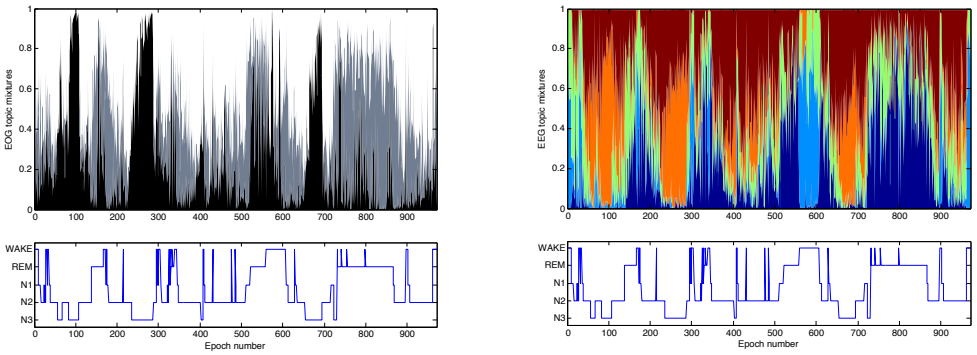
$$\begin{aligned} p\left(\text{NDD} \mid \boldsymbol{f}_{EEG}\right) &= \Big(1+\exp\Big(-0.002+0.363 f_{T_{Amount}}^{Dark\ blue}+0.311 f_{T_{Amount}}^{Orange} \\ &\quad -0.310 f_{T_{Stability}}^{Green}+0.343 f_{T_{Stability}}^{Orange}\Big)\Big)^{-1} \text{ and} \qquad (4.17) \\ p\left(\text{NDD} \mid \boldsymbol{f}_{EOG}\right) &= \Big(1+\exp\Big(-0.003+0.011 f_{Certainty} \\ &\quad +0.320 f_{T_{Amount}}^{Black}\Big)\Big)^{-1}. \qquad (4.18) \end{aligned}$$



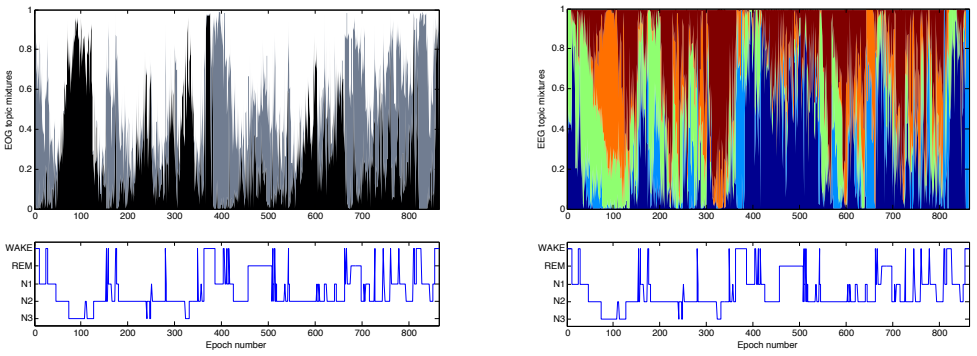
(a) Control subject



(b) PLMD patient



(c) iRBD patient



(d) PD patient

Figure 4.5: Examples of EOG (left) and EEG (right) topic mixture diagrams from a control subject (a), a PLMD patient (b), an iRBD patient (c) and a PD patient (d). A sleep epoch of 30 seconds is represented as a vertical colored bin, where each color represents a topic. The amount of color in each bin illustrates the probability of that topic. Below each topic diagram is provided the manually scored hypnogram for comparison.

The logistic regression model initially including all 13 EEG features ended up including four EEG features, and the regression model initially including nine EOG features ended up using two EOG features. The model initially including all 22 features ended up being the same model as the one including the four EEG features, thereby indicating that the EOG features did not contribute the EEG features in this study.

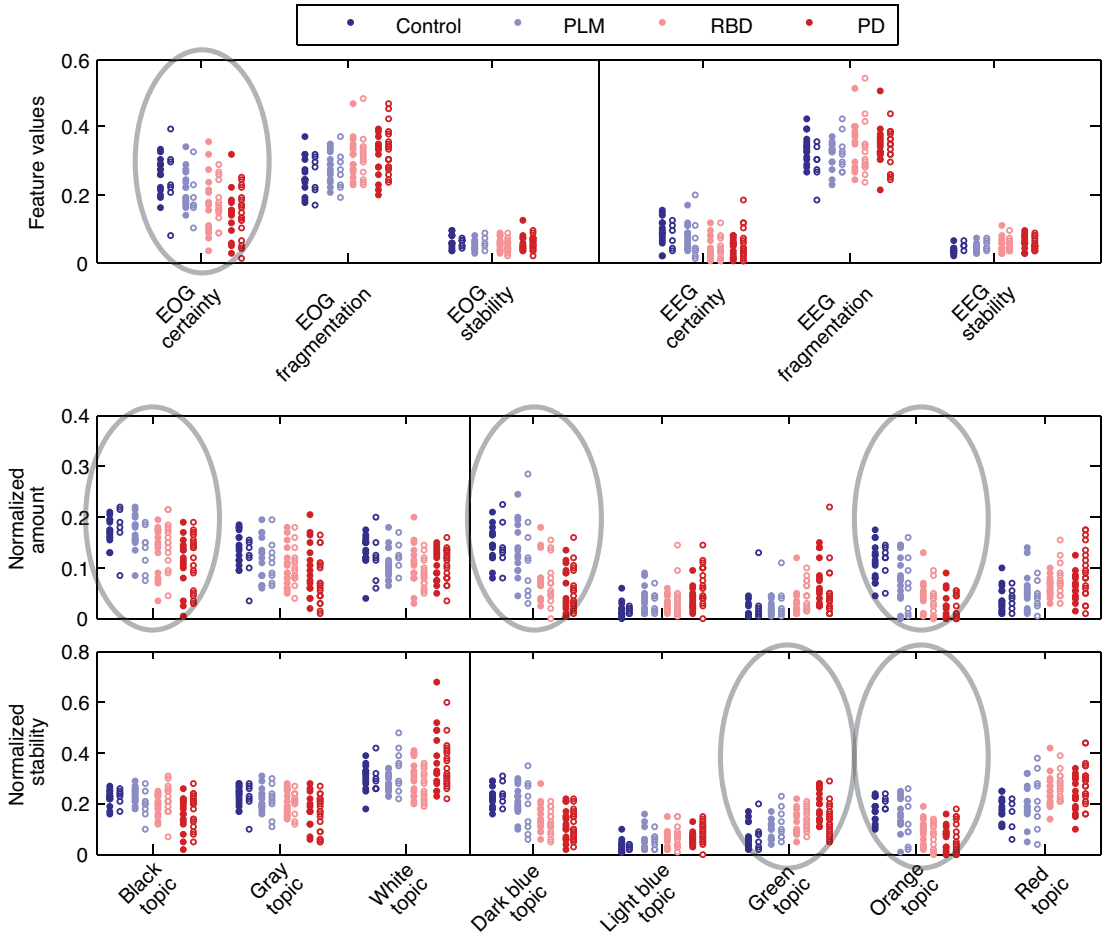


Figure 4.6: Feature values obtained from each subject in the training (filled circles) and validation (open circles) dataset. The EOG features are presented to the left of the vertical black lines and EEG feature are presented to the right of the lines. The larger gray circles indicate the features used in the final regression models.

As seen in equation 4.17, small values of $f_{T_{Amount}}^{Dark\ blue}$, $f_{T_{Amount}}^{Orange}$ and $f_{T_{Stability}}^{Orange}$ and a large value of $f_{T_{Stability}}^{Green}$ increase the probability of belonging to the NDD class. The coefficients show that $f_{T_{Amount}}^{Dark\ blue}$ followed by $f_{T_{Stability}}^{Orange}$, $f_{T_{Amount}}^{Orange}$ and $f_{T_{Stability}}^{Green}$ were the most indicative, implying that the amount of a topic linked to REM sleep and the amount and stability of a topic linked to N3 sleep were considerable higher for the non-NDD class than for the NDD class. The stability of a topic linked to N1/N2 sleep was less for the non-NDD class than for the NDD class.

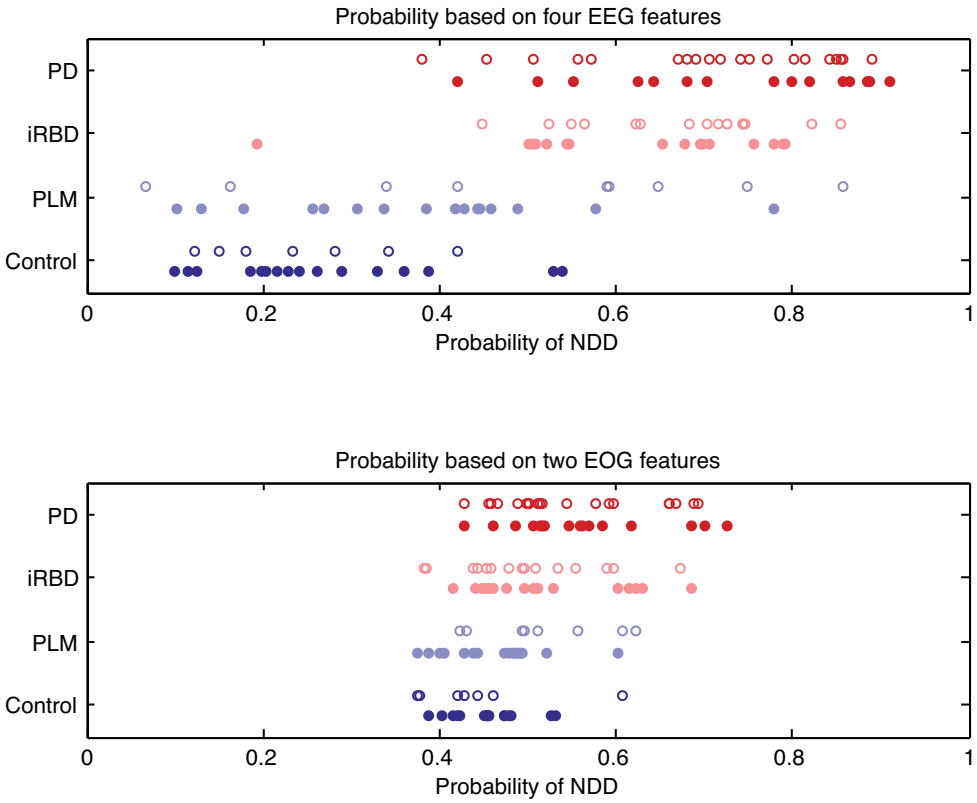


Figure 4.7: Probabilities of belonging to the NDD class as indicated by the best logistic regression models obtained. Each marker is a subject from the training (filled circles) or validation (open circles) dataset. The top figure shows the results from a model using four EEG features and the lower one shows the results from a model using two EOG features. The closer a subject is to 1, the higher probability of belonging to the NDD class.

As seen in equation 4.18, small values of $f_{T_{Amount}}^{Black}$ and $f_{Certainty}$ increased the probability of belonging to the NDD class, where $f_{T_{Amount}}^{Black}$ was the most indicative of the two. This imply that the amount of EEG activity at EOG sites during N3 sleep was considerable higher and the overall EOG certainty (agreement to EOG topics found in control subjects) was marginally higher for the non-NDD class compared to the NDD class.

The final classification was made by assigning subjects to the NDD class when $p(\text{NDD} \mid \boldsymbol{f}) \geq 0.5$. Table 4.5 presents the performance measures for the final regression models and the optimal subset of features chosen for each model. Finally, table 4.6 presents performance measures for different subgroups of the data in order to analyze the effect of age and gender.

As seen in figure 4.7, the distribution of PLMD patients in the validation dataset was relatively different for that of the training dataset. Together with the small amount of control and PLMD patients, this distribution difference is considered the main reason for the small specificity measures obtained. The best model was the one including four EEG features, and the EOG features were found to be non-supportive as they did not add value to the model initially including all features. Although the best model seem to perform better for younger female

Input features		Feature subset	Performance measures	
EOG	EEG		Training dataset	Validation dataset
X	X	1: EEG topic "dark blue" (∼ REM) amount 2: EEG topic "orange" (∼N3) stability 3: EEG topic "orange" (∼N3) amount 4: EEG topic "green" (∼N1/N2) stability	AUC: 93.1% Sens: 93.8% Spec: 87.5%	AUC: 84.3% Sens: 91.4% Spec: 68.8%
X		1: EOG topic "black" (∼ N3) amount 2: EOG overall certainty	AUC: 79.3% Sens: 65.6% Spec: 87.5%	AUC: 64.6% Sens: 57.1% Spec: 62.5%
	X	1: EEG topic "dark blue" (∼REM) amount 2: EEG topic "orange" (∼N3) stability 3: EEG topic "orange" (∼N3) amount 4: EEG topic "green" (∼N1/N2) stability	AUC: 93.1% Sens: 93.8% Spec: 87.5%	AUC: 84.3% Sens: 91.4% Spec: 68.8%

Table 4.5: Performance measures for the three final regression models. The model initially including EEG and EOG features ended up being the same as the model initially including only the EEG features. Sens: Sensitivity; spec: Specificity.

Dataset subgroup	Counts [non-NDD / NDD]		Performance measures	
	Training dataset	Validation dataset	Training dataset	Validation dataset
Subjects aged <60 years	28 (21/7)	18 (9/9)	AUC: 99.3% Sens: 100% Spec: 95.2%	AUC: 99.3% Sens: 88.9% Spec: 77.8%
Subjects aged $\geq$ 60 years	36 (11/25)	33 (7/26)	AUC: 87.3% Sens: 92.0% Spec: 72.7%	AUC: 79.1% Sens: 92.3% Spec: 57.1%
Female subjects	26 (18/8)	19 (10/9)	AUC: 88.9% Sens: 75.0% Spec: 94.4%	AUC: 91.1% Sens: 77.8% Spec: 80.0%
Male subjects	38 (14/24)	32 (6/26)	AUC: 91.7% Sens: 100% Spec: 78.6%	AUC: 71.8% Sens: 96.2% Spec: 50.0%

Table 4.6: Performance measures for different subgroups of the data where the best logistic regression model including four EEG features was used to classify the subjects. Sens: Sensitivity; Spec: Specificity.

subjects, the area under the Receiver Operating Characteristic (ROC) Curve (AUC) measures for the validation dataset in table 4.6 illustrate that the findings in this study were not solely due to age and gender effects.

4.3 Paper IV: Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinon's disease

This study investigates stability of REM sleep, NREM sleep and wake as well as REM-NREM and wake-sleep transitions in control subjects and patients with PLMD, iRBD or PD. The sleep stability and transitions were determined by a data-driven approach as well as by manually labeled stages to investigating whether the topic model approach is beneficial when evaluating iRBD and PD patients. The topic model used is similar to the one described in paper III in section 4.2, but in stead of building one topic model based on EEG and one based on EOG, this approach combine the two modalities to better match the manually scorings. Again, EMG data was left out of the analysis in order to focus on other alterations or potential biomarkers than the affected EMG tonus seen in these patients. The overall methodology of this method is seen in figure 4.8.

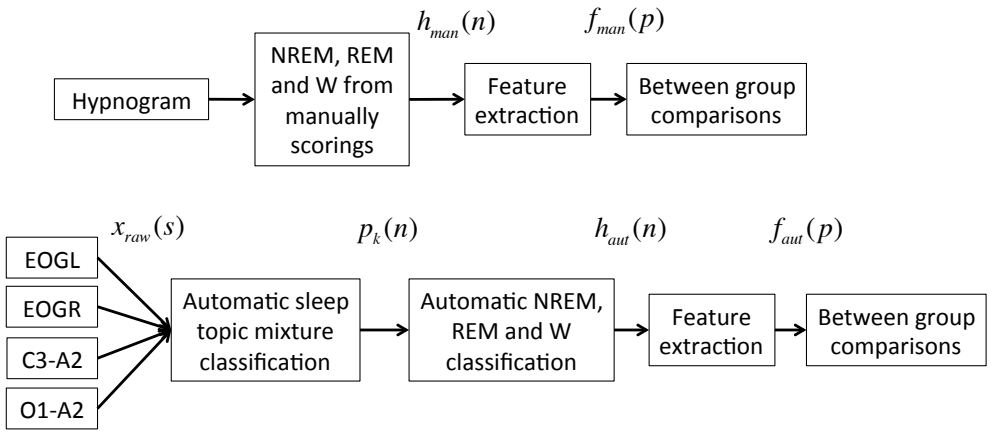


Figure 4.8: The overall methodology of study IV. Features were derived from vectors, h_{man} and h_{aut}, indicating NREM, REM and W stages determined by manually scorings based on the hypnogram or automatic labeled stages, respectively. The automatic labeled stages were found based on topic probability mixtures $p_k(n)$ found from an automatic sleep topic mixture classification model built on EEG and EOG data. Between group comparisons were made between the six groups included in this study.

4.3.1 Methods: Topic modeling for automatic sleep staging

Subjects and recordings

A total of 27 patients with PD, 23 patients with iRBD, 25 patients with PLMD and 23 control subjects aged 40 years or more and with no history of movement disorder, DE behavior or other previously diagnosed sleep disorders were included in this study. As in the other

studies, the patient evaluations included a comprehensive medical and medication history and a PSG analyzed according to the AASM standard. A MSLT was performed in any cases where narcolepsy was suspected. Medication affecting sleep (hypnotics, antidepressants, antipsychotics) was not allowed, except for dopaminergic treatments. The PD patients were divided in those with (PD^+) and those without (PD^-) RBD determined by the presence of RSWA as well as clinical complaints. The iRBD patients were divided in those with a total score of nine or less ($iRBD^-$) and those with a total score of 10 or more ($iRBD^+$) on the RBD Screening Questionnaire (RBDSQ) [80]. This division of patient groups was done in order to investigate if the presence of RBD in PD patients and the self-reported complaints in iRBD patients had an influence on the measures. The cutoff on the RBDSQ score was chosen to divide the iRBD patients into those with major self-reported DE ($iRBD^+$) and those with minor ($iRBD^-$) on a scale of 1-13 where all iRBD patients have to have a score of minimum 5. The PLMD patients were included as an additional control group and did not show any signs of neurodegeneration or RSWA, and could be considered as solely PLMD patients. The quality of the PSG data was individually evaluated and excluded if the quality was unacceptable. Demographic and PSG data for the six groups are summarized in table 4.7.

Automatic labeling of sleep

Neurodegeneration have been reported to cause changes in sleep micro- and macroarchitecture, thereby interfering with sleep scoring. Sleep is especially altered in iRBD and PD patients making it difficult to identify stage-dependent characteristics, and consequently, sleep stage scoring this pathology is associated with high inter-rater variability [21] [39]. In this study, this challenge was solved by letting a data-driven topic model label the sleep epochs. The automatic detector is fully described in the co-authored paper [42], and was developed in a Master's project simultaneously and in a similar way as the ones described in section 4.2 (paper III). The aim for the Master's project was to build a supervised topic model to match manually scorings and a wide range of settings were tried, such as which derivations to include, number of topics, word length and number of symbols to use in the symbolization. The method was optimized on the same PSG data as used in this thesis, and with the aims for that project, the best topic model was found to be build on information from both EEG (the central and frontal derivation) and EOG data (left and right). Also, the number of topics $K = 6$ and a word length of three were found to be the optimal. Figure 4.9 illustrates the development of the sleep topic mixture classification model used in this study.

Each epoch was divided into non-overlapping 1-s segments, yielding a word feature vector $T(m)$ expressed as,

$$T(m) = \begin{bmatrix} X_{C3-A2}(m) \\ X_{O1-A2}(m) \\ T_{EOG}(m) \end{bmatrix} \quad \text{where} \quad X(m) = \begin{bmatrix} x_\delta(m) \\ x_\theta(m) \\ x_\alpha(m) \\ x_\beta(m) \end{bmatrix} \tag{4.19}$$

	Controls	PLMD	iRBD$^-$	iRBD$^+$	PD$^-$	PD$^+$
Total counts (Male/Female)	23(7/16)	25(13/12)	12(9/3)	11(10/1)	8(5/3)	19(13/6)
Age [years, $\mu \pm \sigma$]	56.7±9.2	56.9±11.6	61.8±6.8	66.3±7.2	68.8±8.4	63.7±6.7
BMI $\left[\text{kg/m}^2, \mu \pm \sigma\right]$	23.1±2.5	25.7±3.8	25.1±2.6	25.9±3.6	24.1±3.6	25.9±2.9
SE [$\%, \mu \pm \sigma$]	86.2±10.9	86.1±7.5	82.5±7.4	85.8±6.7	67.9±15.5	80.3±9.9
BT [min, $\mu \pm \sigma$]	484±81.2	431±50.0	518±107.3	474±71.1	427±53.8	461±88.3
W [$\%, \mu \pm \sigma$]	13.8±10.9	13.9±7.5	17.5±7.4	14.2±6.7	32.1±15.5	19.7±9.9
REM [$\%, \mu \pm \sigma$]	18.9±6.6	18.5±6.1	17.7±7.1	16.3±7.1	9.3±5.3	14.6±10.8
N1 [$\%, \mu \pm \sigma$]	8.8±4.6	8.4±7.2	7.5±2.6	10.8±5.4	7.3±6.0	12.6±9.9
N2 [$\%, \mu \pm \sigma$]	44.7±9.4	44.7±10.9	40.0±10.9	43.0±10.3	39.2±12.9	39.3±15.1
N3 [$\%, \mu \pm \sigma$]	13.8±7.3	14.5±9.1	17.3±9.9	15.7±11.2	12.1±9.1	13.9±17.6
LM index [no/hour, $\mu \pm \sigma$]	18.1±13.1	56.5±36.8	49.9±32.0	31.6±22.9	39.2±40.1	42.7±44.4
PLM index [no/hour, $\mu \pm \sigma$]	7.7±7.4	36.5±24.6	14.9±16.3	23.0±17.1	11.2±11.5	10.8±10.3
Disease duration [years, $\mu \pm \sigma$]	NA	NA	5.3±9.8	11.3±12.0	7.7±5.7	4.6±3.3
RBD score [$\mu \pm \sigma$]	NA	NA	7.7±1.6	11.1±1.0	2.5±1.1	9.7±2.3

Table 4.7: Demographic and PSG data for the six groups enrolled in study IV. The PD patients were divided into those with RBD (PD$^+$) and those without (PD$^-$), and the iRBD patients were divided into those with a total score of ≤ 9 (iRBD$^-$) and those with a total score of >9 (iRBD$^+$) on the RBD screening questionnaire. Disease duration is stated as years from clinical diagnosis (PD patients) or self-reported subjective RBD-symptoms (iRBD patients). SE: Sleep efficiency; BT: Time in Bed; LM: Leg movements; PLM: Periodic leg movements.

represents the spectral power in the clinical EEG frequency bands, and where $T_{EOG}(m)$ is the same as stated in equation 4.4. The transformation into EEG words was done as described in section 4.2.1 assigning symbols 1 to 5 to obtain T^w_{EEG} stated in equation 4.5. The transformation into EOG words was done as described in section 3.3.1 assigning symbols 1 to 4 to obtain T^w_{EOG} as stated in equation 3.18. The word length was chosen to be three, thereby yielding a total of $2 \times 4 \times 5^3 = 1000$ possible EEG words (2 derivations, 4 clinical bands, 5^3 word combinations) plus $3 \times 4^3 = 192$ possible EOG words. A fingerprint of each epoch thereby consist of a distribution across 1192 different words holding information from EEG

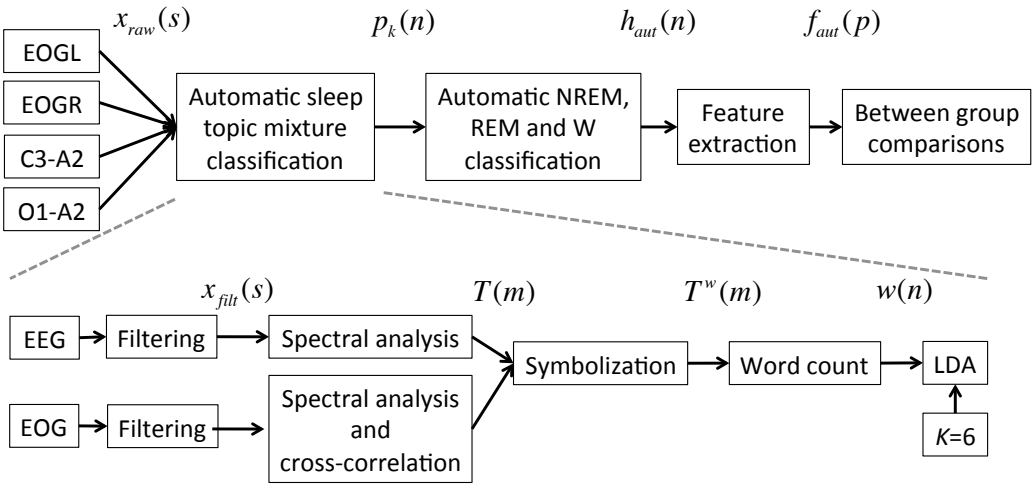


Figure 4.9: Illustration of the conversion of raw EEG and EOG signals, $x_{raw}(s)$, to vectors $\mathbf{w}(n)$ describing the words of each sleep epoch n. The word vectors are given as input to the Latent Dirichlet Allocation (LDA) model, which outputs the probability mixture $p_k(n)$ across the $K = 6$ topics.

and EOG, described by a word vector described as,

$$\mathbf{w}(n) = \begin{bmatrix} \mathbf{w}_{EEG}^{C3-A2}(n) \\ \mathbf{w}_{EEG}^{O1-A2}(n) \\ \mathbf{w}_{EOG}^{ll}(n) \\ \mathbf{w}_{EOG}^{rr}(n) \\ \mathbf{w}_{EOG}^{lr}(n) \end{bmatrix} = \begin{bmatrix} \#\text{"111" in epoch } n \text{ of } X_{C3-A2,\delta}^{w} \\ \vdots \\ \#\text{"555" in epoch } n \text{ of } X_{C3-A2,\beta}^{w} \\ \#\text{"111" in epoch } n \text{ of } X_{O1-A2,\delta}^{w} \\ \vdots \\ \#\text{"555" in epoch } n \text{ of } X_{O1-A2,\beta}^{w} \\ \#\text{"111" in epoch } n \text{ of } X_{ll}^{w} \\ \vdots \\ \#\text{"444" in epoch } n \text{ of } X_{ll}^{w} \\ \#\text{"111" in epoch } n \text{ of } X_{rr}^{w} \\ \vdots \\ \#\text{"444" in epoch } n \text{ of } X_{rr}^{w} \\ \#\text{"111" in epoch } n \text{ of } X_{lr}^{w} \\ \vdots \\ \#\text{"444" in epoch } n \text{ of } X_{lr}^{w} \end{bmatrix}, \tag{4.20}$$

where # denotes the total number. The topic model used in this study was developed and trained using the same datasets as described in section 4.2.1. The main difference between the two papers is that the one described in section 4.2.1 used EEG and EOG separately and did not attempt to match the manually scorings, whereas the topic model used in this study and described in [42] combine the two modalities and try to best match manually scorings

from control subjects. For a more detailed description of the optimization of the topic model used in this study, see paper [42].

For each epoch n, the topic model outputs the posterior probability $p_k(n)$ for each of six sleep topics k. Examples of epochs containing one topic with high certainty are seen in figure 4.10. As we were only interested in REM sleep, NREM sleep and wake in this study, the probabilities for topic 1, 2 and 3 were combined into NREM, and by letting the dominant probability decide, each epoch was labeled REM, NREM or W, described by,

$$h_{aut}(n) = \begin{cases} \text{W if } \arg\max_j (p_j(n)) = 4 \\ \text{REM if } \arg\max_j (p_j(n)) = 5. \\ \text{NREM if } \arg\max_j (p_j(n)) = \{1, 2, 3\} \end{cases} \tag{4.21}$$

By doing so, the overall mean accuracy rates for detecting NREM, REM and W ranged from 70% for patients with PD to 77% for control subjects when comparing with manually single-scored hypnograms. It should be noticed that topic 6 was left out, but this combination of topic probabilities was found to give the highest performance when detecting REM, NREM and W.

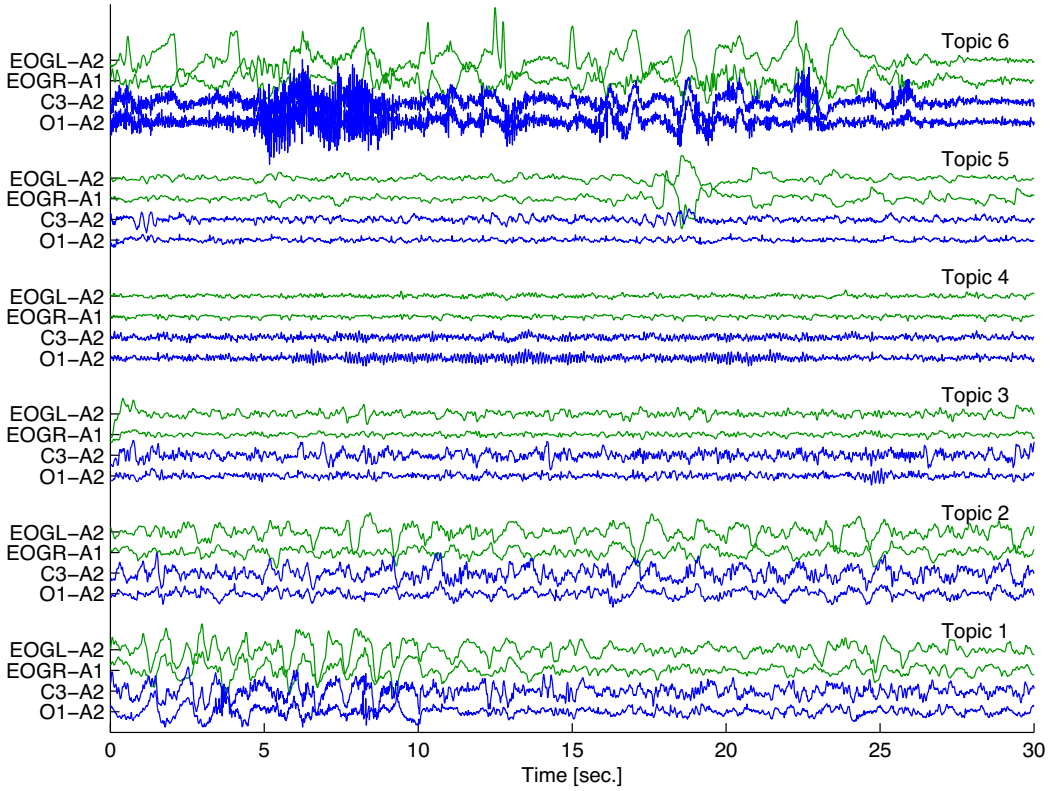


Figure 4.10: Examples of epochs containing one topic with high certainty. Dominating topic certainty in the respective epochs are: Topic 6: $p_6 = 46.8\%$, topic 5: $p_5 = 88.9\%$, topic 4: $p_4 = 68.9\%$, topic 3: $p_3 = 82.9\%$, topic 2: $p_2 = 86.0\%$ and topic 1: $p_1 = 34.7\%$. Figure taken from [42].

Analysis of transitions and stability

Two transition measures and three stability measures were defined and analyzed in this study. The measures were computed based on the automatic labeled epochs, described by $h_{aut}(n)$ as well as the manually scored epochs. Wake-sleep transitions were defined as the number of shifts from any sleep stage (N1, N2, N3, REM) to wakefulness or *vise versa*, and REM-NREM shifts were defined as the number of transitions between any NREM stage (N1, N2, N3) to REM sleep or *vice versa*. Stability measures were computed as the number of passages from 1) REM to REM, 2) NREM to NREM or 3) W to W. The transition measures were defined as the frequency per minute of total time in bed, and the stability measures were defined as the frequency of passages between two REM, NREM or W epochs per minute of the total time spent in these stages, respectfully. The features are summarized in table 4.8 and a vector $f_{aut}(p)$ hold the ones derived from the automatic labeled sleep stages while $f_{man}(p)$ hold the ones derived from the manually labeled sleep stages.

Feature name	Feature explanation	Total no per subject	
		Automatic scorings	Manual scorings
REM-NREM transitions	Number of transitions from NREM to REM or REM to NREM per minute of total time in bed	1	1
Wake-sleep transitions	Number of transitions from sleep to wake or wake to sleep per minute of total time in bed	1	1
REM stability	Number of toggles from REM to REM per minute of total time spent in REM sleep	1	1
NREM stability	Number of toggles from NREM to NREM sleep per minute of total time spent in NREM sleep	1	1
W stability	Number of toggles from W to W per minute of time spent in W	1	1

Table 4.8: Overview of the feature groups computed in study IV.

To test for between-group differences, Wilcoxon rank-sum tests were performed on the two transition and three stability measures, each having 15 between-group comparisons thereby summing up to $5 \times 15 = 75$ tests in total. A significance level of $p < 0.05$ was used and the Benjamini-Hochberg procedure was used to correct for multiple testing using a false discovery rate (FDR) at level $q = 0.10$. This procedure was done for the measures computed based on the manual and automatic labeled sleep epochs.

4.3.2 Results

Figure 4.11 presents the frequency of the sleep-wake and REM-NREM shifts and the stability measures for REM sleep, NREM sleep and wake measured by the manually scored as well as the automatic scored sleep epochs. Only the results for the comparisons that remained significant after FDR correction are presented.

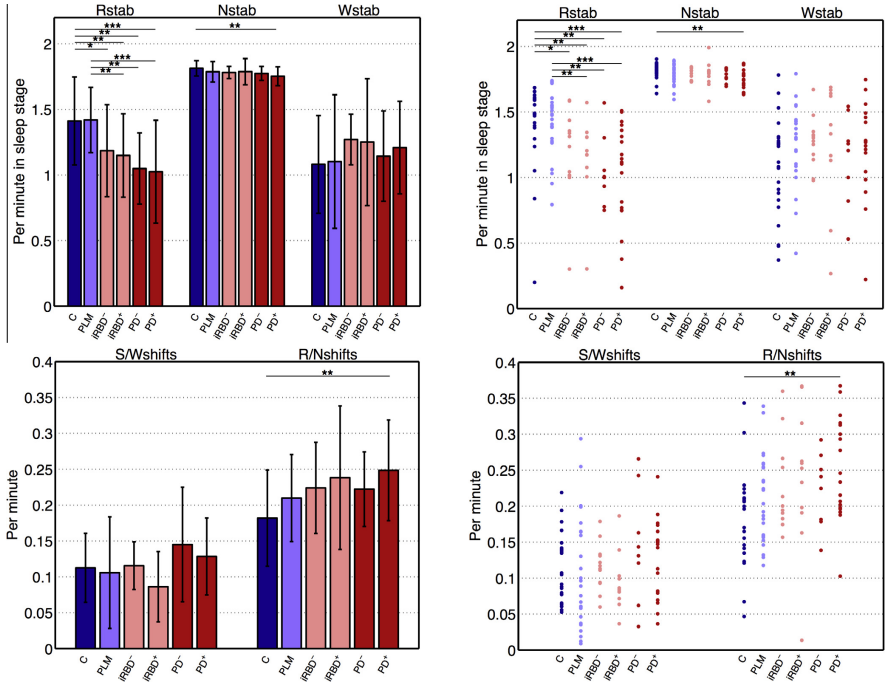
It was found that both groups of iRBD patients showed significantly lower REM stability compared to control subjects, and that the iRBD$^+$ group also showed significantly lower REM stability compared with PLMD patients. Both groups of PD patients showed lower REM stability compared to controls and PLMD patients, and the PD$^+$ group also showed significantly lower NREM stability than controls and significantly more REM-NREM shifts than controls. Finally, non-significant trends were seen for the iRBD$^-$ group that showed a trend of lower REM stability compared with PLMD patients, and the iRBD$^+$ group that showed a trend of more REM-NREM shifts than controls. No significant between-group differences were found between any of the iRBD and PD patients, although the PD patients showed a trend towards lower REM stability compared with iRBD patients. Finally, no significant differences were found for any of the measures computed from the manually scored sleep epochs.

Comparing the results obtained from the two different scoring techniques, it is seen that the stability measures computed from the manually scored hypnograms are all greater than the measures computed from the automatic labeled sleep epochs. This is especially true for the REM stability measure, which is considered the main reason for why the REM-NREM sleep transitions are greater for the automatic staging techniques compared to the manually scorings.

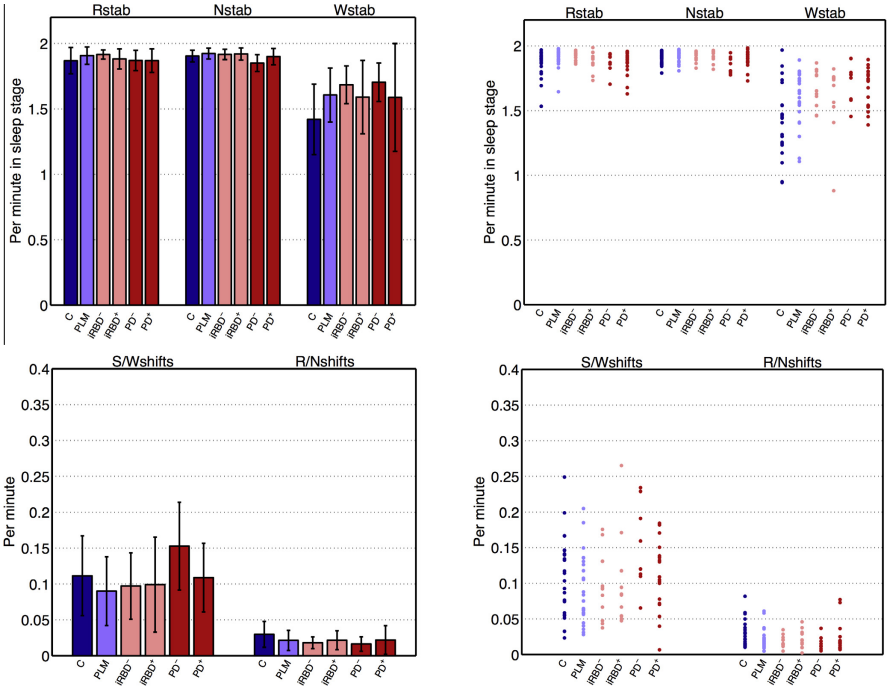
4.4 Conclusive remarks

Study III:
In the first study (Paper III) several features reflecting characteristics of EEG and EOG topics were analyzed separately and combined. The distribution and stability of EEG topics linked to REM and N3 sleep were found to be the most indicative characteristics. The study demonstrates that characteristics derived from EEG topics were better suitable for classifying iRBD and PD patients compared to characteristics derived from EOG topics. The EOG characteristics were found not to be supportive for EEG, maybe because the EOG topic diagrams reflect the same sleep structures as the EEG topic diagrams, but to a less complex and varied degree. However, the topic diagrams as well as the features derived from both EOG and EEG showed differences between iRBD/PD and controls/PLMD patients.

Study IV:
In the second study (Paper IV), the topic model approach was used to determine sleep stability and transitions and the data-driven method was found to be supportive when evaluating iRBD and PD patients. Specifically, it was found that the ability to maintain NREM sleep and in specific REM sleep is progressively affected in iRBD and PD, probably reflecting the successive involvement of brainstem areas from early on in the disease. The manually scorings failed to identify these alterations, maybe due to poor visual identification of micro-sleep structures needed to terminate REM sleep.



(a) Automatic scored epochs



(b) Manually scored epochs

Figure 4.11: (a): The frequency of wake-sleep and REM-NREM shifts and the stability of REM sleep, NREM sleep and wake for the six groups based on the automatic labeled epochs. (b): The frequency of wake-sleep and REM-NREM shifts and the stability of REM sleep, NREM sleep and wake for the six groups based on the manually scored epochs. The measures are illustrated by bars (group mean $\pm$ one standard deviation) to the left and as dots (one per subject) to the right. Only the results for the comparisons that remained significant after false discovery rate (q = 0.10) correction are illustrated. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Overall:
The two studies illustrate that iRBD and PD patients suffer from micro-sleep fragmentation expressed both in EEG and EOG, which interferes with and challenges today's way of scoring sleep. As the REM sleep specifically was found to hold indicative information and as the topic models did not include EMG activity, the studies illustrate that the neurons regulating the cortical activation during REM sleep are affected by the neurodegeneration. Finally, the two studies suggest iRBD to be an intermediate stage between controls and PD patients, consistent with Braak's hypothesis [6].

It is concluded, that LDA is a useful approach for analyzing sleep. The proposed methods proved to reflect the overall sleep pattern and gave a more detailed representation of sleep compared to the AASM standard. Allowing sleep epochs to be expressed as mixtures of stages revealed that the transitions between sleep stages are continuous rather than abrupt. The continuous processes and the association between topics should be investigated in much more detail in future studies. Also, it should be emphasized that although the studies illustrate EEG and EOG alterations in patients, the characteristic of the alterations are not identified. The complexity of LDA makes it feasible of detecting concurrent topics that are associated to only few words, but it also makes it difficult to investigate the direct link between words and topics.

Overall, this research area conclude:

- that topic models developed on EOG and/or EEG data from control subjects can be useful when evaluating EOG and EEG from patients with PLMD, iRBD or PD.
- that sleep data from patients with iRBD or PD do not fit well into topics derived from control subjects, illustrating that both EOG and EEG in these patients are altered.
- that features reflecting sleep stage transitions and stability as well as amount of time spent in each sleep stage can be extracted from the topic diagrams and thereby be used as an automated indication of alterations in iRBD/PD patients compared to controls/PLMD patients.

5

Sleep spindles

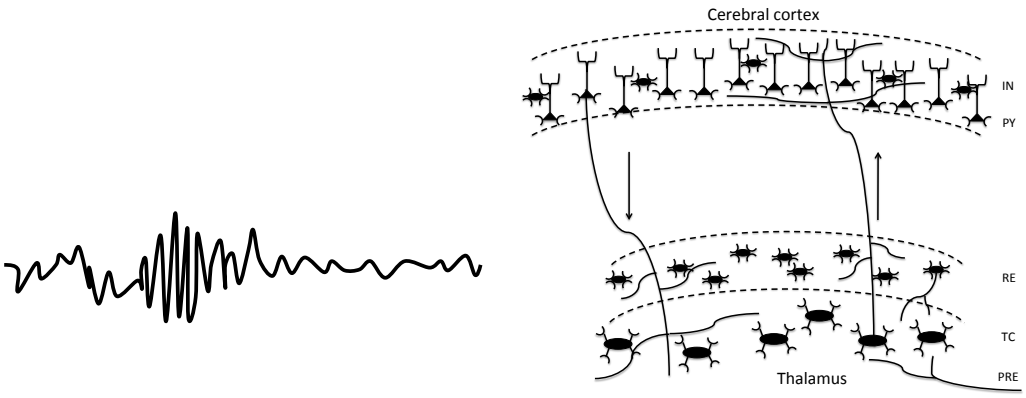
Objective *Sleep spindles (SS) are controlled by a complex interaction involving thalamic, limbic and cortical areas. Thalamus hold the primary role in generating and controlling spindles, and as it is located between the cerebral cortex and the midbrain, it is sensitive to involvement in neurodegenerative disorders (NDD). This chapter is composed upon paper V and VI, and serves to investigate the biomarker potential of SS. In Paper V the spindle density is investigated as a potential biomarker of PD, whereas paper VI serves to identify spindle alterations in PD patients.*

5.1 Background

Sleep spindles (SS) are EEG hallmarks of NREM sleep, and are mostly present during N2 sleep. They are generated in thalamus, as evidenced by the fact that thalamectomy eliminates SS in the sleep EEG [43]. Many neurological connections exist between the thalamus and the cerebral cortex and these are thought to play an important role in the generation and controlling of different micro-sleep structures, including SS. In figure 5.1 is seen a typical SS and an illustration of the generation of SS.

The generation of SS begins in the di-synaptic circuit between thalamic reticular neurons (RE) and thalamocortical relay cells (TC), where spontaneously spindle-like oscillations are generated. The oscillations are conveyed to the cortex by the TC cells, and cortical pyramidal cells (PY) mediate a cortico-thalamic feedback as they project back to the RE and TC cells, that conveyed the spindle. In the cortex, cortical interneurons (IN) ensure cortico-cortical connections, which together with the cortico-thalamic feedback loop are important in synchronizing the occurrence of SS over widespread thalamic and cortical areas [43].

The functional meaning as well as the triggers of SS are not fully known, but as the TC cells also receive input from prethalamic fibers (PRE) arising from specific sensory systems as well as systems located in the brainstem and posterior hypothalamus [79] [78], SS are thought to have a sleep-preserving role through inhibition of sensory input [30]. Also, they are assumed to play an important role in memory consolidation during sleep, synaptic plasticity and cognition [27] [28] [46] [69] [79]. The formation of SS begins in the infant brain [30], and



(a) (b)

Figure 5.1: (a): A typical sleep spindle. (b): Illustration of thalamic and cortical cells and neural interactions involved in the generation and controlling of sleep spindles. PRE: Prethalamic fibers; TC: thalamocortical pyramidal cells; RE: thalamic reticular neurons; PY: cortical pyramidal cells; IN: cortical interneurons.

as several decreasing trends in SS characteristics occur with age [30] [58], they are suggested to have an important role in normal cognitive function.

According to the AASM standard, SS have an oscillating frequency of 11-16 Hz and a duration of 0.5-3 seconds [34]. They are nearly sinusoidal waves with a characteristic Gaussian amplitude envelope and an ability to stand out from the background EEG. Previously, the frequency profile has been stated at 12-14 Hz or 12-16 Hz, and in some studies SS are referred to as two different kinds - the slow ones with frequencies of 11.5-14 Hz and the fast ones with frequencies of 14-16 Hz [43]. As the manual scoring relies on subjective assessment, it can be a difficult and tedious task to identify SS, and the inter-scorer agreement rate for scoring SS in normal sleep has been indicated to be as low as $70\pm8\%$ [87]. Also, automated SS detectors have been indicated to show significant variance between them when identifying SS in normal sleep [84].

SS identification and characterization in pathological sleep is not well studied, but has been indicated to be altered in patients with PD and other NDDs such as Dementia, Alzheimer's disease and mild cognitive impairment [46] [63] [85]. Generally, the SS density has been reported to be low in these patient groups, and although not specific to a certain patient group, they could be useful as biomarkers of disease progression or therapeutic efficacy [48] [54] [56].

Besides age and pathology, terms such as memory consolidation, pharmacological interventions and pre-PSG conditions have been reported to affect SS activity. Taken together with the different definitions of SS, the reliance on subjective evaluation and the high inter-expert variability, the impact of the different terms are very difficult to state.

5.1.1 Research hypothesis

The main hypothesis for this research area was that SS hold potential for being a PD biomarker. The research area also serves to illustrate the challenges faced when developing automated spindle detectors to be used on pathological sleep.

The aim with this research area was to:

- develop an automated SS detector suitable for detecting SS with an acceptable performance.
- investigate whether the presumed reduced SS activity is apparent in iRBD patient as well as in PD patients.
- identify changes in SS density and morphological characteristics in patients with PD compared to controls.
- identify specific SS features that may be useful as prognostic biomarker of PD by relating them to disease duration and cognitive measures.

It should be mentioned, that the development of the SS detector (first- and second-mentioned aim) was carried out in the Master's project preceding current PhD dissertation. Inclusion of more data, re-evaluation of results, re-writing, submission and revision of the manuscript was, however, carried out during the PhD, and Paper V is thus included in this research area.

5.2 Paper V: Decreased sleep spindle density in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease

In this study an automated SS detector was developed, and used to measure the SS density in control subjects, patients with iRBD and patients with PD with or without RBD. The overall methodology of the development of the SS detector is seen in figure 5.2.

5.2.1 Methods: Development of an automated spindle detector

Subjects and recordings

A total of 15 control subjects, 15 patients with iRBD, 15 patients with PD and RBD and 15 patients with PD without RBD were enrolled in this study. The inclusion and exclusion criteria for controls and patients were as stated in the previous studies, and the demographic data for the four groups are summarized in table 5.1.

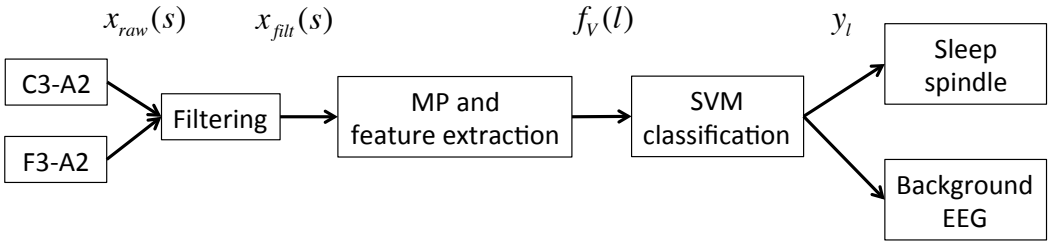


Figure 5.2: Schematic illustration of the development of the SS detector described in paper V. Data from the central and frontal EEG derivation was bandpass filtered and decomposed by Matching Pursuit (MP) from where features were extracted. The feature vector $f_V(l)$ holds 12 features per segment l. A Support Vector Machine (SVM) was used to classify each segment of data as "sleep spindle" or "background EEG".

	Controls	iRBD	PD-RBD	PD+RBD
Frequency (Male/female)	15(6/9)	15(12/3)	15(8/7)	15(11/4)
Age [years]	58.3±9.5	60.1±7.4	61.9±6.1	62.4±5.2
BMI $[\mathrm{kg/m^2}]$	23.2±2.8	24.4±3.1	24.7±2.2	26.0±3.2
SE [%]	88.9±8.4	85.6±8.3	82.8±7.9	85.4±9.7
BT [min]	480±47.5	489±95.3	443±67.2	445±71.8
Wake [no(%)]	1606(11)	2220(15)	2387(18)	1889(14)
REM [no(%)]	271(19)	2893(20)	1808(13)	1761(13)
N1 [no(%)]	1205(8)	1238(8)	1191(9)	1623(12)
N2 [no(%)]	6491(45)	5909(40)	5817(44)	5957(45)
N3 [no(%)]	2388(17)	2423(17)	2097(16)	2128 (16)

Table 5.1: Demographic data for the four groups included in study V. PD-RBD denotes patients with PD without RBD and PD+RBD denotes patients with PD and RBD. SE: Sleep efficiency; BMI: Bode mass index; BT: Time in bed.

Feature extraction

The C3-A2 and F3-A2 EEG derivations, both sampled with a sampling frequency f_s=256 Hz, were used to extract appropriate features for detecting SS. In this study, the Matching Pursuit (MP) method was used for feature extraction, and after the signals were band-pass filtered from 2 to 35 Hz, they were decomposed using MP. In MP, a given signal is represented by a weighted sum of Gabor atoms, $g_\gamma(t)$, expressed as [52],

$$g_\gamma(t) = K(\gamma)e^{-\pi\left(\frac{t-u}{s}\right)^2}\cos\left(\omega\left(t-u\right)+\phi\right) \tag{5.1}$$

where $\gamma = \{u, s, \omega, \phi\}$ represents time shift u and width s in seconds, frequency ω in rad/s, phase ϕ in rad and $K(\gamma)$ represents a normalization scaling factor. The decomposition is an iterative process, where the input signal for each step is matched with Gabor atoms and where the Gabor atom with the highest correlation is selected and subtracted the input signal. The iterative process in MP is illustrated in figure 5.3.

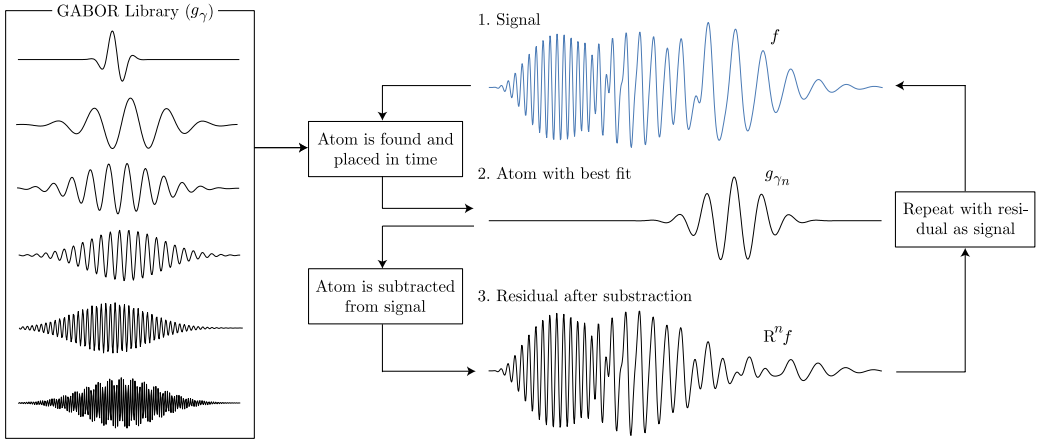


Figure 5.3: Schematic illustration of the iterative process in Matching Pursuit (MP). The Gabor atoms in the dictionary to the left are compared to the analyzed signal, and the one with the biggest inner product with the signal is chosen. The atom is adjusted and subtracted from the analyzed signal and the resulting residual becomes the new signal to be decomposed. The process stops when a the residual is below a given threshold. Figure taken from [77].

The process stops when the residual is below a given threshold, and the original signal $f(t)$ can be expressed as [52],

$$f(t) = \sum_{n=0}^{M-1} \left\langle R^n f(t), g_{\gamma_n}(t) \right\rangle g_{\gamma_n}(t) + R^M f(t) \tag{5.2}$$

where $\left\langle R^n f(t), g_{\gamma_n}(t) \right\rangle$ represents the inner product of the nth atom and the signal $R^n f(t)$ and $R^M f(t)$ denotes the residual signal after approximating $f(t)$ by using M Gabor atoms. The time-frequency distribution of the signal energy is derived by adding Wigner-Ville distributions of selected atoms, yielding [52]

$$\begin{aligned} WV_f(t,\omega) \quad = \quad & \sum_{n=0}^{M-1} \left| \left\langle R^n f(t), g_{\gamma_n}(t) \right\rangle \right|^2 WV_{g_{\gamma_n}}(t,\omega) + \\ & \sum_{n=0}^{M-1} \sum_{k=1, k\neq n}^{M-1} \left\langle R^n f(t), g_{\gamma_n}(t) \right\rangle \left\langle R^k f(t), g_{\gamma_k}(t) \right\rangle WV_{g_{\gamma_n}, \gamma_k}(t,\omega), \end{aligned} \tag{5.3}$$

where WV_f and $WV_{g_{\gamma_n}}$ indicate the Wigner-Ville distribution of the signal f and the Gabor atom g_{γ_n}, respectively, and $\langle , \rangle$ represents the inner product. Finally, the energy density of the signal $f(t)$ is found by removing the cross-terms of the Wigner-Ville transform, yielding

$$E_f(t,\omega) = \sum_{n=0}^{M-1} \left| \left\langle R^n f(t), g_{\gamma_n}(t) \right\rangle \right|^2 WV_{g_{\gamma_n}}(t,\omega). \tag{5.4}$$

The features were all computed from the energy densities obtained from signal windows of 2 s with a 1-s overlap. For each 2-s window l and for each EEG derivation, six features were computed, yielding 12 features in total described as,

$$f_V(l) = \begin{bmatrix} f_V^{C3-A2} \\ f_V^{F3-A2} \end{bmatrix} \quad \text{where} \quad f_V^d = \begin{bmatrix} E_l^d(t,\omega) \text{ for } \omega < 11 \cdot 2\pi \\ E_l^d(t,\omega) \text{ for } 11 \cdot 2\pi \leq \omega \leq 16 \cdot 2\pi \\ E_l^d(t,\omega) \text{ for } 16 \cdot 2\pi < \omega \\ \log\left(|\left\langle R^i l^d(t), g_{\gamma_i}(t)\right\rangle|^2 WV_{g_{\gamma_i}}(t,\omega)\right) \\ \log\left(\max_{t,\omega}\left(E_l^d(t,\omega)\right)\right) \\ \omega \text{ at } \max_{t,\omega}\left(E_l^d(t,\omega)\right) \end{bmatrix} \tag{5.5}$$

where i denotes the first Gabor atom with a frequency of 11 Hz $\leq f \leq$ 16 Hz, $l^d(t)$ denotes the signal segment and E_l^d denotes the energy density found by equation 5.4 for the EEG derivation d. An overview of the features is provided in table 5.2. Before given as input to SVM, the 12 features were normalized with respect to the 95th percentile of the feature values across all 2-s windows for that specific subject.

Feature name	Feature explanation	Total no per 2-s window	
		F3-A2	C3-A2
Energy features	Energy in the frequency bands f <11 Hz, 11 Hz $\leq f \leq$ 16 Hz and $f >$ 16 Hz computed by equation 5.4.	3	3
Gabor atom features	The logarithm of the energy contribution of the first Gabor atom with a frequency of 11 Hz $\leq f \leq$ 16 Hz.	1	1
Max energy density point features	The logarithm of the maximum energy point in the energy density found by equation 5.4 and the corresponding frequency	2	2

Table 5.2: Overview of the features computed in study V. For each 2-s window, six features were computed for each EEG derivation. Three of the features reflected energy in the frequency bands below, in and above the spindle frequency band, one reflected the energy contribution from the first Gabor atom with a frequency in the spindle band, and two reflected the energy and the frequency of the maximum energy point in the energy distribution.

Classification

In this study the Support Vector Machine (SVM) algorithm was used to classify the SS [18]. SVM is a binary supervised learning method, where two classes are separated by a separating hyperplane in a high-dimensional feature space [18]. The hyperplane is found based on a labeled training dataset described as $\{x_i, y_i\}_{i=1}^L$, $x_i \in \Re^D$ where $y_i \in \{-1, 1\}$ indicates which class the sample vector x_i belongs to. The samples that impact the slope of the hyperplane

the most are called support vectors and they all satisfy the constraint,

$$\left.\begin{array}{lll} \langle \mathbf{x}_i, \mathbf{w} \rangle + b & \geq 1 - \xi_i & \text{for} \quad y_i = 1 \\ \langle \mathbf{x}_i, \mathbf{w} \rangle + b & \leq -1 - \xi_i & \text{for} \quad y_i = -1 \end{array}\right\} \Rightarrow \; y_i \left(\langle \mathbf{x}_i, \mathbf{w} \rangle + b \right) - 1 + \xi_i \geq 0 \quad \forall i \tag{5.6}$$

where $\mathbf{w}$ is the normal to the hyperplane, b is a shifting constant and $\xi_i \geq 0 \, \forall i$ is a slack variable included to relax the constraints of the fully separable case by introducing a penalty to misclassified samples. The separating hyperplane is thus found by solving the problem summarized to [18],

$$\begin{cases} \min \left(\dfrac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{L} \xi_i \right) \\ y_i \left(\langle \mathbf{w}, \mathbf{x}_i \rangle + b \right) - 1 + \xi_i \geq 0 \quad \forall i \\ \xi_i \geq 0 \quad \forall i \end{cases} \tag{5.7}$$

where C is a user-defined cost parameter indicating the penalty for misclassification. The problem is solved by introducing Lagrange multipliers, and the final SVM classifier is defined through $\mathbf{w}$ and b describing the optimal orientation of the separating hyperplane. Classification is carried out by evaluating on which side of the separating hyperplane a new sample point $\mathbf{x}'$ lies, or simply by evaluating the sign of the function,

$$h\left(\mathbf{x}'\right) = \langle \mathbf{x}', \mathbf{w} \rangle + b. \tag{5.8}$$

In cases where the classes are not linearly separable, a kernel $K\left(x_i, x_j\right)$ is used to map the data into a Euclidean space H where the classes can be linearly separated. In this study a Radial basis function kernel $K\left(\mathbf{x}_i, \mathbf{x}_j\right) = e^{\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2} \right)}$ was used where the cost parameter C and the kernel-specific parameter $\gamma = \dfrac{1}{2\sigma^2}$ were optimized by grid search.

The SS detector was trained using a number of randomly selected sleep epochs from 13 of the control subjects. All SS in these epochs were manually labeled where only the F3-A2, C3-A2 and O1-A2 EEG derivations were visible for the SS scorer. The spindle criteria stated in the AASM standard were used, and each second of data was labeled either SS (1) or background EEG (-1). Table 5.3 summarize the data used for training the automated SS detector. A total of 882 SS were manually labeled.

Sleep stage	W	REM	N1	N2	N3	Total
Number of epochs (%)	0 (0)	4 (1)	13 (4)	330 (88)	28 (7)	375 (100)

Table 5.3: Distribution of the different sleep stages for use in the development of the SS detector.

The SVM classifier was trained using the leave-one-subject out approach, and the overall performance measures were calculated as the mean across the 13 runs.

5.2.2 Results

In figure 5.4 is seen the Receiver Operating Characteristic (ROC) curve for the final SS detector. The area under the ROC curve (AUC) reached 91.0% based on the leave-one-subject-out strategy. Choosing the best point on the curve as the point ranking the false and true positive rate equally, the mean sensitivity reached 84.7% and the mean specificity reached 84.5%. These were considered appropriate for the purpose of this study.

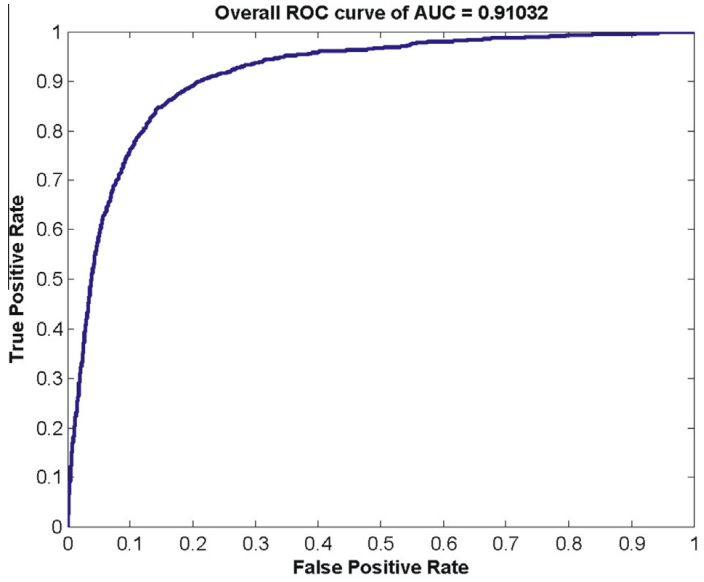


Figure 5.4: The overall Receiver Operating Characteristic (ROC) curve with a mean area under the curve (AUC) measure of 91.0% based on the leave-one-subject-out strategy used for automatic spindle detection in paper V.

The SS detector was applied on all data between lights off and lights on for the 4×15 subjects included in this study. SS density was defined as SS/min and was found for N1, N2, N3, all NREM and REM sleep. Table 5.4 summarizes the mean values and standard deviations of SS densities obtained for the four groups.

	N1	N2	N3	All NREM	REM
Controls	4.4±1.6	6.2±1.5	5.6±1.3	6.0±1.3	2.2±1.4
iRBD	4.4±1.7	4.7±1.9	4.1±2.4	4.5±1.8	2.8±1.4
PD-RBD	4.4±1.7	5.1±1.8	4.9±2.3	5.0±1.5	2.4±1.4
PD+RBD	4.4±2.1	4.2±1.9	3.6±2.1	4.2±1.8	3.6±2.2

Table 5.4: Means and standard deviations of the SS densities for the four groups obtained using the SS detector. SS density was defined as SS/min.

Unpaired two-sample t-tests were performed to establish between-group differences. When comparing the control group with a group of patients, the t-tests were one-sided in order to

test whether the control group had a significantly higher mean SS density compared to the patients. Comparing two patient groups, the t-tests were two-sided to establish whether the means differed from each other. Figure 5.5 illustrates the significant differences found.

It was found that patients with iRBD, and patients with PD with or without RBD had significantly lower SS densities compared to controls in N2 and in all NREM sleep combined. Also, patients with iRBD and patients with PD and RBD showed significantly lower SS densities compared to controls in N3 sleep. No significant differences were found between any two groups of patients, and no significant differences were found in REM sleep.

In this study, no performance values were obtainable for the SS detector used on pathological sleep. Also, the detector labeled each second (limited by the resolution of the manual scorings), which is considered to be low when detecting SS.

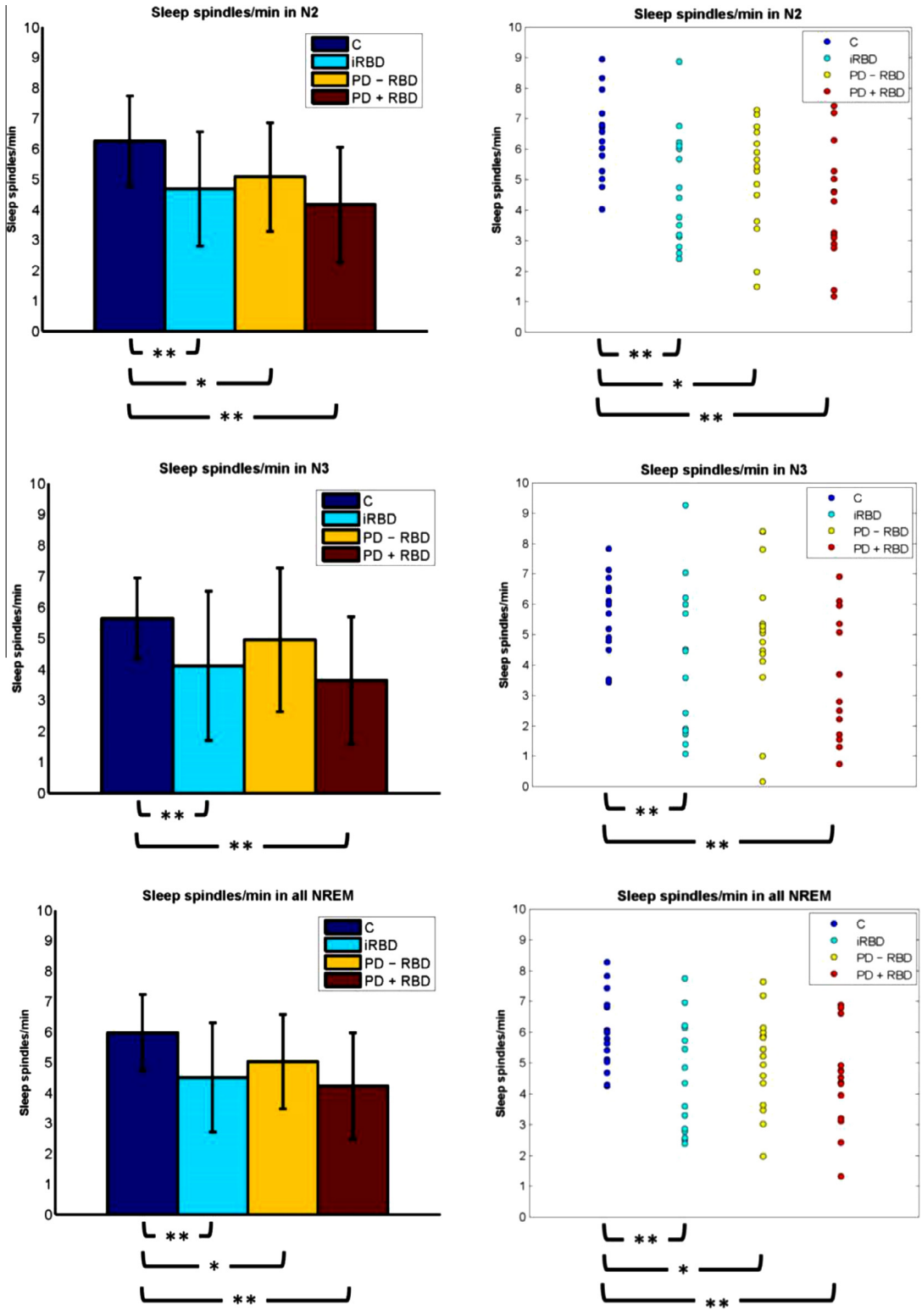


Figure 5.5: Sleep spindle (SS) densities obtained for the four groups in N2, N3 and all NREM sleep combined. The group means $\pm$ one standard deviation are illustrated as the heights of bars with error bars (left) and the individual SS densities are illustrated as dots (right). *: p<0.05; **: p<0.01.

5.3 Paper VI: Sleep spindle alterations in patients with Parkinson's disease

In this study, five independent sleep experts manually labeled and rated SS from control data and data from PD patients by assigning a confidence score for each SS. For SS meeting a group consensus rule, features describing morphological characteristics were computed and between-group comparisons were performed in order to determine if SS are altered in patients with PD, and if they can serve as prognostic biomarker of disease. It was decided to obtain SS characteristics from manually scored rather than automatically identified SS to allow spindles not bounded to the limits given by the AASM standard. Also, a SS detector developed based on normal sleep only might be biased and hinder true alterations to be identified. Due to these reasons and the fact that the developed SS detector described in paper V had a resolution of one second, it was decided that morphology measures of spindles should be extracted from manually labeled spindles rather than automatically detected.

The overall methodology of this study is seen in figure 5.6.

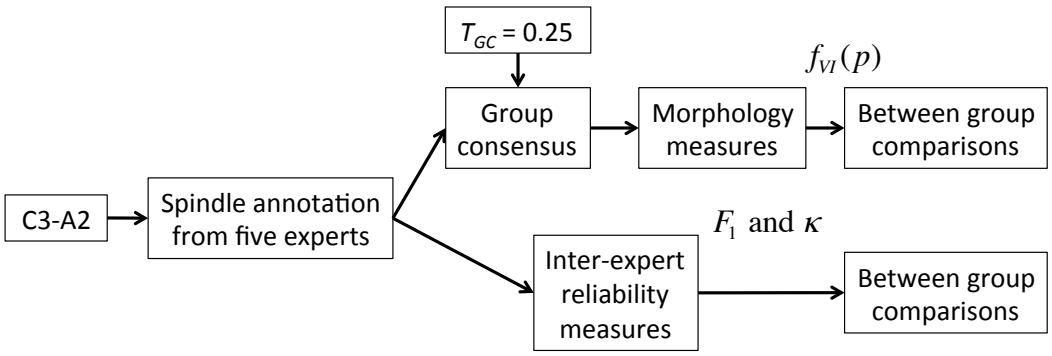


Figure 5.6: Overall methodology of paper VI. Five sleep experts identified spindles (SS) in data from the central EEG derivation. Applying a threshold T_{GC}=0.25 on the average score across samples, a group consensus of SS was obtained, from which morphology measures were extracted. Also, the inter-expert reliability measures, F_1-score and Cohen's κ, were computed and between group comparisons were conducted for the morphology and inter-expert reliability measures.

5.3.1 Methods: Generation of a spindle database and computation of morphology measures

Subjects and recordings

A total of 15 PD patients and 15 sex- and age-matched control subjects with no history of movement disorder, dream-enacting (DE) behavior or other previously diagnosed sleep disorders were included in this study. As in the previous studies, patients were evaluated

with a comprehensive medical and medication history and a PSG analyzed according to the AASM standard [34]. None of the patients had Dementia at inclusion, but one of them later developed Multiple System Atrophy (MSA) of the Parkinsonian type (MSA-P). Medications known to affect sleep were not accepted, although dopaminergic treatments were permitted. The quality of each PSG recording was individually examined and accepted. Demographic data and PSG variables for the two groups are seen in table 5.5.

Characteristics	PD patients	Controls	P
Total counts (Male/female)	15 (7/8)	15 (7/8)	-
Age [years]	62.7±5.8	62.9±5.9	0.90
BMI $\left[\mathrm{kg/m^2}\right]$	25.3±3.5	22.1±2.5	0.02
Disease duration [years]	6.7±4.5	NA	-
Hoehn & Yahr stage	2.0±1.2	NA	-
UPDRS part III "on"	20.9±7.0	NA	-
ACE	90.2±4.8	NA	-
Levodopa equivalent dosage [mg]	621.1±301.5	NA	-
Levodopa use [n (%)]	10 (67)	NA	-
Dopamine agonist use [n (%)]	14 (93)	NA	-
Sleep efficiency [%]	79.7±14.1	87.1±8.4	0.09
Time in bed [min]	448.1±82.0	499.6±63.7	0.07
LM index [no/hour]	31.8±34.8	30.4±35.3	0.91

Table 5.5: Demographic and PSG data for the two groups included in study VI. BMI: Body mass index; UPDRS: Unified Parkinson's disease rating scale; ACE: Addenbrooke's cognitive examination; LM: Leg movements.

Manual labeling of spindles

The inter-rater variability of scoring SS has been reported to be as low as 70±8% [87], which clarify the need for several scorers when building a reliable dataset of SS. For each subject, eight blocks of five consecutive epochs of N2 sleep were randomly selected in between lights off and lights on. In these blocks, five independent sleep experts identified SS in the C3-A2 EEG derivation, which was the only one visible. The scoring was performed in a Matlab-based program "EEG viewer", which mimics a standard sleep scoring program in a clinical setting. Before scoring SS, the C3-A2 signal was filtered in the program with a notch filter at 50 Hz and a band-pass filter with cutoff frequencies at 0.3 and 35 Hz, as indicated by the AASM standard.

For each SS, the experts assigned a confidence score of 1 for "definitely SS", 0.75 for "probably a SS" and 0.5 for "maybe a SS". Background EEG was labeled with 0. The final SS identification used for morphology measures were defined using a group consensus rule where the confidence scores were averaged at each sample point and aggregated into a single consensus. Sample points with an average score of higher than $T_{GC} = 0.25$ were included

in the final group consensus. An illustration of the group consensus rule is seen in figure 5.7. This scoring procedure allows the experts to label doubtful spindles, and if identified by several experts, they are included in the final spindle database. In this way, the database will hold spindles that have had an ability to stand out from the background EEG but not necessarily bound to the limits given by AASM.

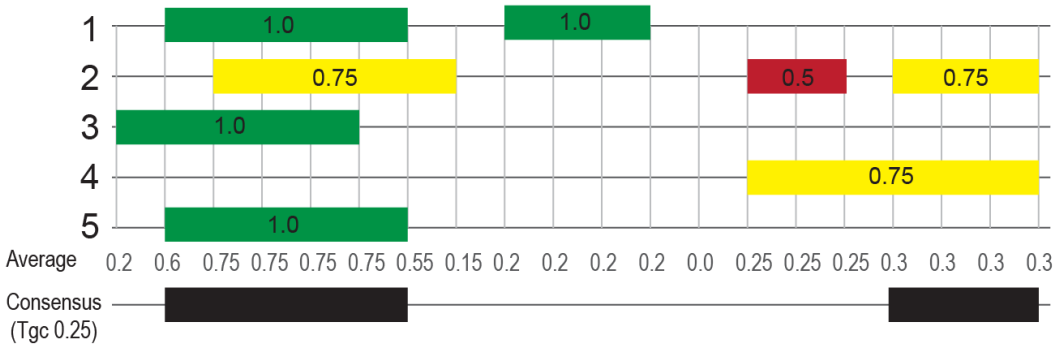


Figure 5.7: Illustration of the group consensus rule applied to obtain the final spindle identification. Taken from [84].

Feature extraction

For the spindles included in the group consensus, five features reflecting morphology measures were computed. These include 1) the duration dur in seconds, 2) the oscillation frequency f_{osc} in Hz, 3) the maximum peak-to-peak amplitude in mV of the original spindle signal (A_{p2p}) and of a high-pass filtered signal (A^{hp}_{p2p}), 5) a symmetry measure sym in % and 6) the density in spindles/min. Table 5.6 summarizes the morphology features computed.

Feature name	Feature explanation	Total no	
		Per spindle	Per person
Duration	Duration in seconds.	1	-
Oscillation frequency	Frequency in Hz computed based on the number of extrema points and the duration	1	-
Max peak-to-peak amplitude	Amplitude in μV computed as the max vertical distance between two extrema points	2	-
Symmetry	Percentage of samples before max peak-to-peak amplitude	1	-
Density	Number of spindles per minute	-	1

Table 5.6: Overview of the features computed in study VI. For each spindle, the duration, oscillation frequency, max peak-to-peak amplitude and symmetry measure were computed. The max peak-to-peak amplitude was computed twice for each spindle; once for the spindles filtrated from 0.3-35 Hz (A_{p2p}), and once where frequencies below 4 Hz were removed (A^{hp}_{p2p}). The density measure was computed for each subject.

Letting f_s = 256 Hz be the sampling frequency, K the number of extrema points, A_e a vector holding the original amplitude values of the extrema points and A_e^{hp} a vector holding the amplitude values of the high-pass filtered spindle signal, the morphology features were defined as,

$$f_{VI} = \begin{cases} dur = \dfrac{\#\text{ samples}}{f_s} \\ f_{osc} = \dfrac{K}{2 \cdot dur} \\ A_{p2p} = \max\left(\mid A_e\left(k+1\right) - A_e\left(k\right) \mid\right), k = 1, 2, ..., K-1 \\ A_{p2p}^{hp} = \max\left(\mid A_e^{hp}\left(k+1\right) - A_e^{hp}\left(k\right) \mid\right), k = 1, 2, ..., K-1 \\ sym = \dfrac{\#\text{ samples before point of} A_{p2p}}{\#\text{ samples}} \\ density = \dfrac{2 \cdot \#\text{ SS}}{\#\text{ epochs reviewed}} \end{cases} . \tag{5.9}$$

The extrema points were detected using matlab's *findpeaks*-function on a 5-point moving average smoothed version of the SS signal and with a minimum peak-to-peak distance of 11 samples. The extrema points were defined as the sum of the maxima points found by applying the *findpeaks*-function directly, and the minima points found by applying it on the flipped spindle signal. These settings were chosen as they were considered best for estimating f_{osc} when visually investigating numerous randomly selected examples of SS.

The reason for why two maximum peak-to-peak amplitudes were computed (A_{p2p} and A_{p2p}^{hp}), was to investigate the influence on SS from the high amplitude, low-frequency K-complexes or delta waves. Forward and reverse filtration with a 10th order high-pass filter with a cut-off frequency (-3dB) at 4 Hz was considered suitable for removing low frequency, high amplitude waves that may interfere with the peak-to-peak calculation.

The morphology measures were computed for the spindle identifications for each expert as well as for each spindle in the group consensus. Due to the shortening effect of the group consensus, a duration threshold of dur_{th} = 0.2 s was used on spindles in the group consensus resulting in the exclusion of three spindles. For each measure, two-sided Wilcoxon rank sum tests with a significance level of α = 0.05 were used to test for significance between controls and PD patients.

Inter-expert reliability when scoring spindles

Five independent sleep experts identified SS in the same data, given ten available expert-pair, for which inter-expert reliability measures indicated as the F_1-score and the Cohen's Kappa coefficient (κ) were computed. True positives (TP) and true negatives (TN) define the number of samples where both experts have marked SS and not marked SS, respectively, and false positive (FP) and false negative (FN) define the number of samples where the reference-expert has not marked SS, but the other expert has, and number of samples where the reference-expert

has marked a SS, but the other has not, respectfully. Given these, the inter-expert reliability measures are expressed by,

$$F_1\text{-score} = \frac{2 \cdot R \cdot P}{R + P} \text{ and } \kappa = \frac{\frac{TP+TN}{N} - Pr}{1 - Pr}, \tag{5.10}$$

where $N = TP + TN + FP + FN$ defines the total number of samples and where precision P, recall R and hypothetical probability of change agreement Pr are given as,

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN} \quad \text{and} \tag{5.11}$$

$$Pr = \frac{TP + FN}{N}\frac{TP + FP}{N} + \left(1 - \frac{TP + FN}{N}\right)\left(1 - \frac{TP + FP}{N}\right). \tag{5.12}$$

The F_1-score is the harmonic mean of P and R and reaches its best value at 1 (perfect agreement) and the worst at 0 (no agreement). The κ is often used as an inter-annotator reliability measure, as it takes the agreement occurring by chance into account. It reaches its best value at 1 (perfect agreement), its worst at -1 (no agreement), and when κ = 0, the accuracy is equal to what is expected by chance. κ is often categorized with the labels "poor" (κ < 0.00), "slight" ($0.00 \leq \kappa \leq 0.20$), "fair" ($0.21 \leq \kappa \leq 0.40$), "moderate" ($0.41 \leq \kappa \leq 0.60$), "substantial" ($0.61 \leq \kappa \leq 0.80$) or "almost perfect" ($0.81 \leq \kappa \leq 1.00$) [45].

5.3.2 Results

In figure 5.8 is seen the five morphology measures computed. The maximum peak-to-peak amplitude provided is the one computed from the spindle signal before removal of frequencies below 4 Hz. Table 5.7 summarizes the means and standard deviations for the spindle characteristics found for each of the experts' identifications as well as for the spindles included in the group consensus.

It was found that patients with PD show altered SS compared to controls, as they differed significantly in terms of density, duration, oscillation frequency and maximum peak-to-peak amplitude. Specifically, patients with PD had a lower SS density and showed spindles with a longer duration, lower oscillating frequency and higher max peak-to-peak amplitude both before and after removal of frequencies below 4 Hz.

Both measures of the max peak-to-peak amplitude were significant different between groups for all five experts, the duration and oscillation frequency was significant different between groups for 4/5 experts, and density was significantly different between groups for 3/5 experts. One of the PD patients had a very high spindle density, and considering that this patient later developed MSA, the spindles from this patient were excluded and the results were reanalyzed. Again, same spindle measures were found to be significant, as seen in table 5.8.
In figure 5.9 and 5.10 is seen the morphology measures for the patients when they are sorted according to the disease duration (5.9(a)), Addenbrookse's cognitive examination (ACE) score (5.9(b)), Hoehn and Yahr (H&Y) stage (5.10(a)) and Unified Parkinson's Disease rating

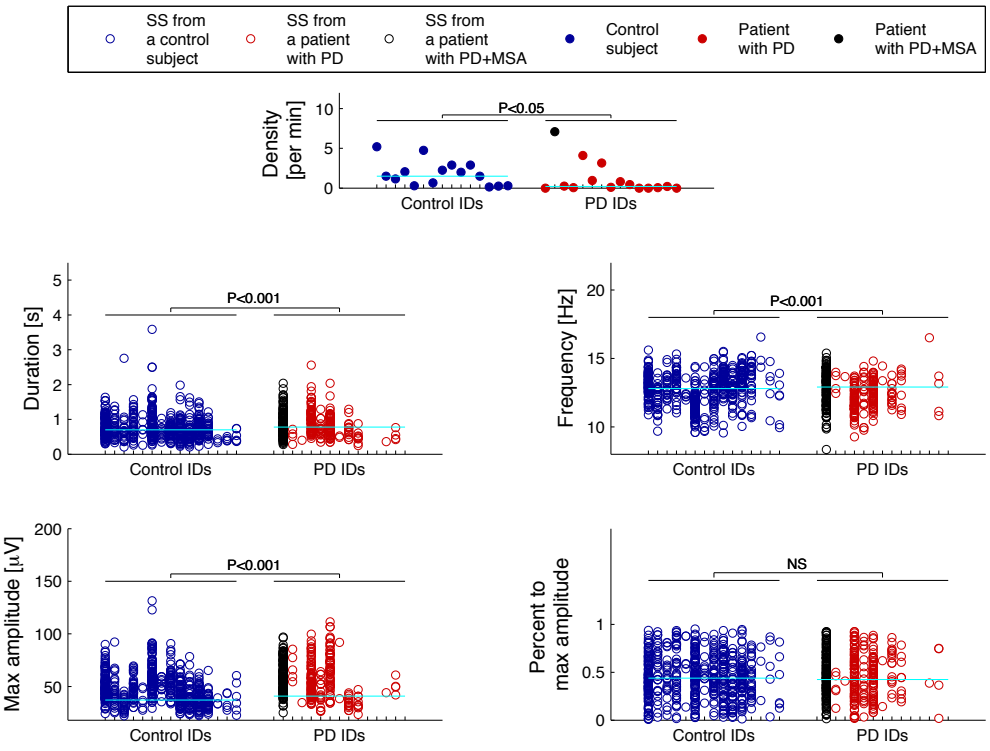


Figure 5.8: Illustration of the five morphological measures computed for the spindles in the group consensus. The maximum amplitude is the one computed from the spindles before removal of frequencies below 4 Hz. The P values were obtained from two-sided Wilcoxon rank sum tests between the two groups.

		dur	f_{osc}	A_{p2p}	A_{p2p}^{hp}	Sym	$Density$
Expert 1	PD	0.93±0.44	12.38±1.27	57.73±17.23	53.87±15.99	0.49±0.23	1.29±2.44
(947 SS)	C	0.84±0.41	12.96±1.27	48.26±15.37	45.42±14.17	0.47±0.24	1.87±1.56
Expert 2	PD	0.66±0.29	12.92±1.24	56.96±18.14	54.38±16.85	0.46±0.23	0.91±1.36
(752 SS)	C	0.67±0.29	13.07±1.11	46.88±15.96	44.16±14.89	0.46±0.23	1.60±1.27
Expert 3	PD	0.74±0.29	12.45±1.22	57.75±17.15	54.89±16.34	0.46±0.24	1.16±1.95
(952 SS)	C	0.68±0.27	12.62±1.34	48.52±15.47	45.43±14.18	0.46±0.23	2.01±1.82
Expert 4	PD	0.88±0.20	12.73±1.14	64.60±16.68	62.40±16.64	0.46±0.22	0.30±0.51
(282 SS)	C	0.77±0.24	13.13±1.00	49.95±14.04	47.51±13.13	0.46±0.21	0.64±0.84
Expert 5	PD	1.19±0.52	11.69±1.24	51.44±18.14	46.20±16.62	0.45±0.25	2.91±2.52
(2135 SS)	C	1.12±0.51	12.03±1.28	45.02±15.73	40.15±13.92	0.45±0.25	4.21±2.14
GC	PD	0.86±0.35	12.51±1.21	57.64±17.34	54.78±16.24	0.47±0.23	1.15±2.06
(901 SS)	C	0.77±0.36	12.80±1.23	48.19±15.55	45.29±14.41	0.46±0.23	1.86±1.57
P		<0.002 ^{1,3,4,5,GC}	<0.02 ^{1,3,4,5,GC}	<0.001 ^{1,2,3,4,5,GC}	<0.001 ^{1,2,3,4,5,GC}	NS	<0.05 ^{1,2,3,GC}

Table 5.7: Mean and standard deviation for the spindle characteristics for each of the experts' identifications as well as for spindles in the group consensus (GC). ^{1}: Significant for expert 1; ^{2}: Significant for expert 2; ^{3}: Significant for expert 3; ^{4}: Significant for expert 4; ^{5}: Significant for expert 5; ^{GC}: Significant for the group consensus. The data investigated is the same across all experts.

		dur	f_{osc}	A_{p2p}	A_{p2p}^{hp}	Sym	$Density$
GC	PD(-MSA)	0.86±0.35	12.27±1.07	56.35±18.97	53.42±17.84	0.47±0.23	0.72±1.28
(759 SS)	C	0.77±0.36	12.80±1.23	48.19±15.55	45.29±14.41	0.46±0.23	1.86±1.57
P		<0.001	<0.001	<0.001	<0.001	NS	<0.007

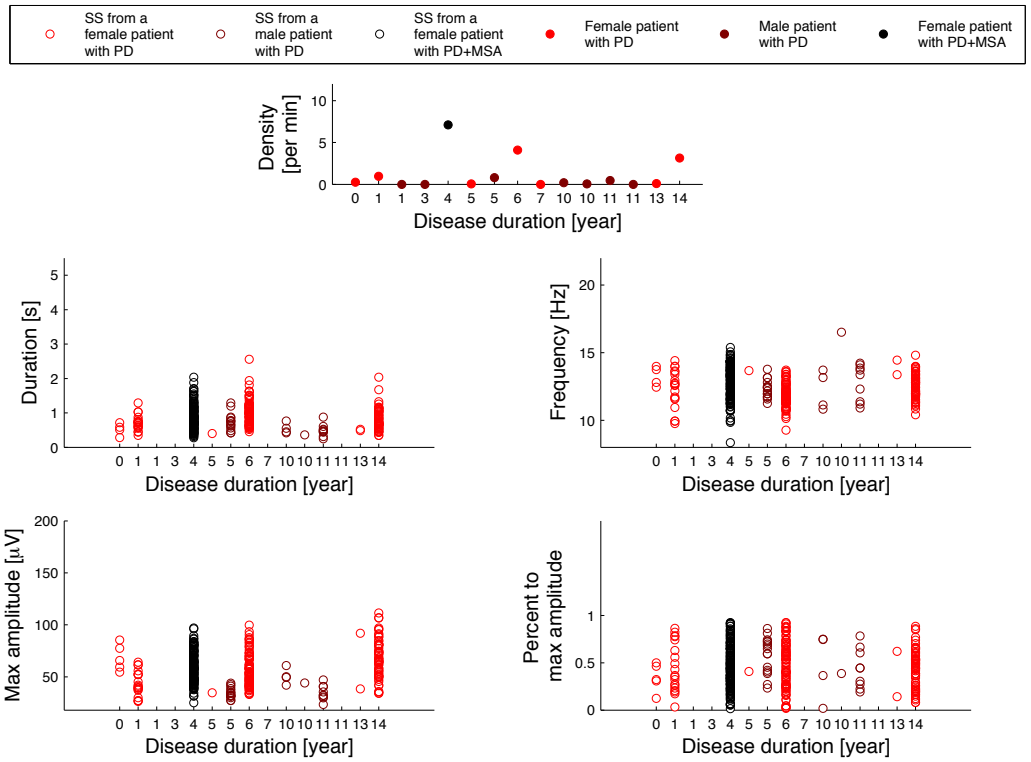
Table 5.8: Mean and standard deviation for the spindle characteristics found for spindles in the group consensus (GC) when spindles from the patients that later developed MSA (PD(-MSA)) were excluded.

scale (UPDRS) part III "on" score (5.10(b)), respectfully. No clear visual trends were seen for any of the disease or cognitive indicators, reflecting that SS can not be used alone as an indicator for disease severity. It may also be that this study includes too few subjects to show a trend due to the great inter-subject variance in SS characteristics, or simply that the cognitive assessment (ACE and UPDRS part III) are too brief.

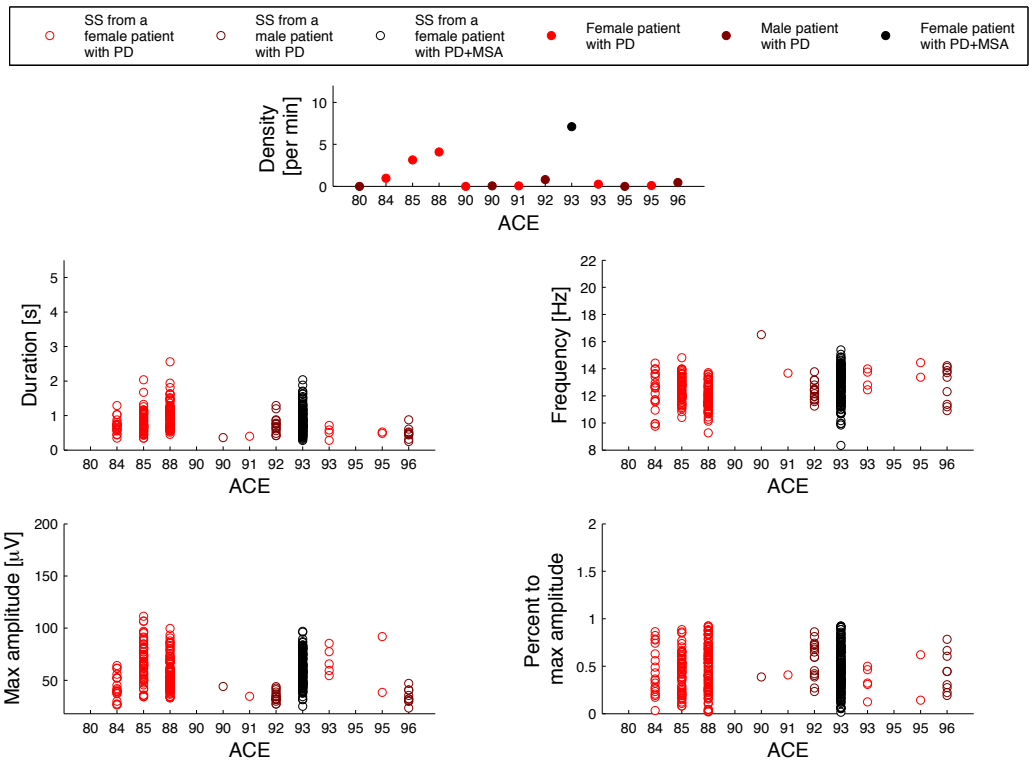
In table 5.9 is seen the fraction of SS included in the group consensus that do not strictly pass AASM criteria for a spindle. A fourth (25.3%) of all spindles in the group consensus did not meet AASM criteria, mostly because they were too short (16.9%) or had an oscillation frequency that were too slow (9.7%). Controls were found to show significantly more spindles not meeting the AASM criteria, i.e. more spindles with a too short duration compared with PD patients. Excluding spindles from the outlier patient (MSA-P), no significant differences were found, but may be due to lack of enough data. When performing between-group comparisons only including SS meeting the AASM criteria, however, the same characteristics were found to be significant between controls and PD patients (results not shown here).

AASM criteria	Total SS (901)	PD SS (344)	PD(-MSA) SS (202)	Control SS (557)	P PD vs. C	P PD(-MSA) vs. C
Duration too short (<0.5 s)	0.169	0.128	0.134	0.194	0.010	NS
Duration too long (>3 s)	0.001	0	0	0.002	NS	NS
Oscillation frequency too slow (<11 Hz s)	0.097	0.90	0.099	0.101	NS	NS
Oscillation frequency too high (>16 Hz)	0.002	0.003	0.005	0.002	NS	NS
At least one criteria not met	0.253	0.212	0.228	0.278	0.027	NS

Table 5.9: Percent of spindles included in the group consensus that do not strictly meet AASM criteria. PD indicate all patients with PD. PD(-MSA) indicate all PD patients except the one that later developed Multiple System Atrophy (MSA).

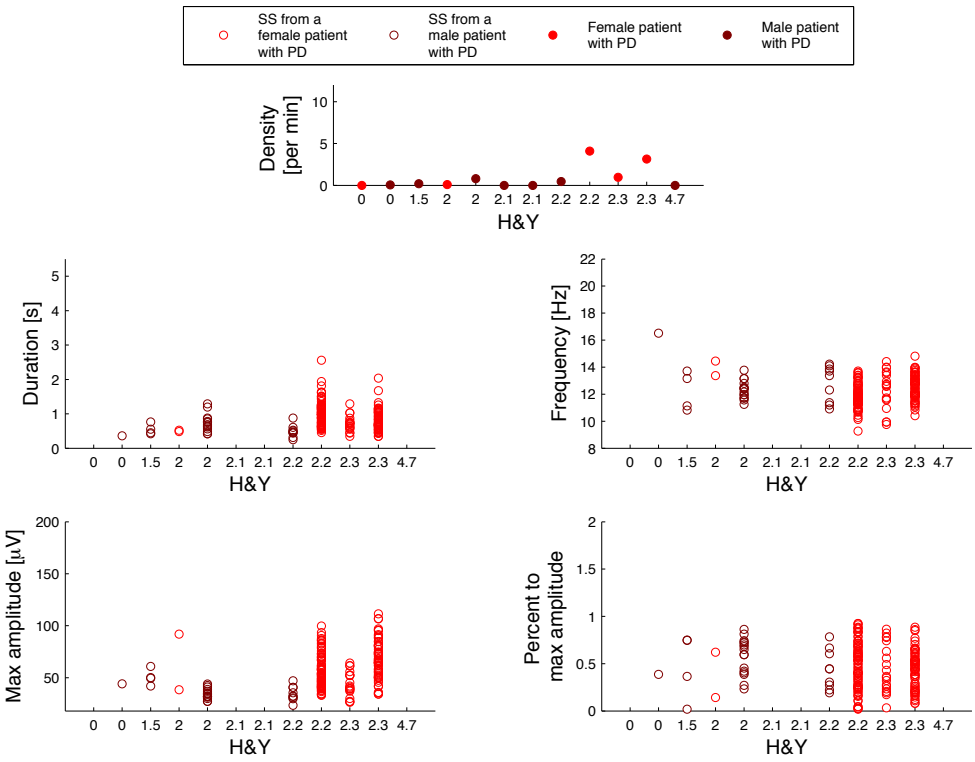


(a)

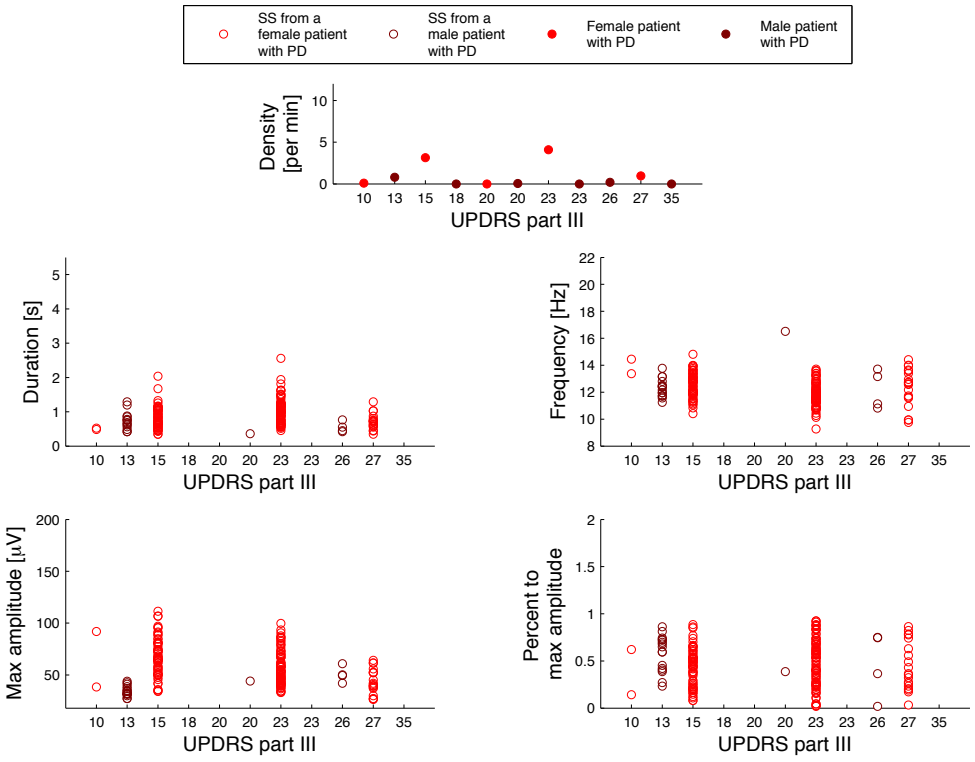


(b)

Figure 5.9: (a): Spindle characteristics for all 15 PD patients sorted according to their disease duration. (b): Spindle characteristics for 13/15 PD patients sorted according to their Addenbrookse's cognitive examination (ACE) score. The max peak-to-peak amplitudes are the ones computed from the spindles before removal of frequencies below 4 Hz.



(a)



(b)

Figure 5.10: Spindle characteristics for 11/15 PD patients sorted according to their Hoehn and Yahr (H&Y) stage ((a)) or their Unified Parkinson's Disease rating scale (UPDRS) score ((b)). The max peak-to-peak amplitudes are the ones computed from the spindles before removal of frequencies below 4 Hz.

Lastly, table 5.10 summarizes inter-expert reliabilities for SS identification, when they were grouped according to their confidence score. In all cases, the agreement rates were lower for scoring SS in PD patients compared to controls, but only the reliability of identifying "definite SS" computed by κ was found to be significantly lower. Also, a trend was seen for a lower κ for "probable/definite SS" in patients compared with controls.

SS group definition	F_1-score		κ		P
	PD	C	PD	C	
Low confidence "maybe"	0.12±0.11	0.17±0.12	0.14±0.11 "slight"	0.16±0.12 "slight"	NS
Medium confidence "probably"	0.13±0.10	0.19±0.11	0.15±0.10 "slight"	0.18±0.11 "slight"	NS
High confidence "definitely"	0.24±0.13	0.32±0.13	0.21±0.13 "fair"	0.32±0.13 "fair"	$<0.005^{\kappa}$
Medium or high confidence "probably/definitely"	0.34±0.15	0.39±0.17	0.28±0.15 "fair"	0.39±0.17 "moderate"	NS

Table 5.10: Mean and standard deviation for the inter-expert reliability measures F_1-score and κ for identifying SS. The mean and standard deviations are taken across the ten expert-pairs available.

5.4 Conclusive remarks

Study V:

In the first study (paper V), an automated SS detector was developed based on features extracted from the MP algorithm and a SVM used to classify each second into "SS" or "background EEG". By use of the detector, this study demonstrated that PD patients, both with and without RBD, as well as iRBD patients show significantly decreased SS activity. The study argues for the problem that the detector is solely trained on control data and that performance values when applying it on data from patients were unobtainable. Also, it argues for the problem that due to the resolution of the manual scorings, the SS can only be detected on a second-by-second basis, which is considered to be too low. Nevertheless, the study illustrates that SS density as determined by an automated detector identifying "control" SS could be a PD biomarker as it was found to be decreased in patients with iRBD or PD.

Study VI:

In the second study (Paper VI), specific morphological characteristics of SS were computed and compared between control subjects and PD patients. Specifically, it was found that the SS density was lower, duration was longer, oscillation frequency slower and max peak-to-peak amplitude higher in PD patients compared to controls. Based on a group consensus of five individual experts' identification of SS, this study demonstrated that spindles are significantly altered in PD patients. The study lacked to demonstrate correlation between any spindle

characteristics and disease duration, severity or cognitive identifiers, and concluded that due to high inter-subject variability in disease progression and severity, longitudinal studies are needed to investigate the clinical utility as well as the prognostic biomarker potential of SS alterations.

Overall:
The two studies demonstrate that SS activity in form of density and the morphological measures duration, oscillation frequency and maximum peak-to-peak amplitude all have potential to be a biomarker of PD.

Although none of the morphology measures were specific nor sensitive enough as an independent biomarker, certain spindle characteristics might have utility as markers of disease progression. SS are according to the AASM standard defined to have a certain duration and frequency profile, but the dilemma is that these criteria are set based on how SS look in normal sleep. Pathological influence can alter the SS to such a degree that they no longer obey the definitions. The two studies illustrate and discuss the challenges faced when developing a SS detector for use in pathological sleep, the high inter-scorer variability in SS identification and the high inter-subject variability in SS activity.

Overall, this research area conclude:

- that a SS detector developed based on features extracted from the MP algorithm and a classification based on a SVM is suitable when investigating SS density, but that it due to a low resolution (second-by-second basis) is not suitable when investigating SS morphology.
- that SS density is decreased in patients with iRBD and patients with PD with or without RBD determined by an automated SS detector.
- that SS density and morphology is altered in patients with PD determined based on a group consensus of five sleep experts' manual SS identifications.
- that no SS morphology alterations have clear correlates with disease duration or cognitive measures, and that much more data on SS, cognitive function and disease severity is needed to investigate such relations.

6

Discussion and conclusive remarks

The overall purpose of this PhD dissertation was to identify new potential PD biomarkers by employing appropriate biomedical signal processing algorithms on sleep signals. Overall, the main findings of this project can be summarized by following statements:

- Patients with iRBD and patients with PD reflect abnormal form and/or timely distribution of EMs during sleep compared to control subjects.
- Patients with iRBD and patients with PD show altered distribution and stability of N3 and REM sleep determined by a data-driven approach compared to control and PLMD patients.
- Patients with iRBD or PD show less stable REM sleep determined by a data-driven approach regardless of the RBDSQ total score and the presence of RBD compared to controls and PLMD patients.
- Patients with PD and RBD show lower NREM stability and more REM-NREM shifts determined by a data-driven approach compared with controls.
- Patients with iRBD and patients with PD with and without RBD show a decreased spindle density determined by an automated SS detector compared to controls.
- Patients with PD show altered spindles with longer duration, slower oscillation frequency and higher max peak-to-peak amplitude compared with controls.

In summary, the findings of this project emphasize that patients with iRBD or PD suffer from micro-sleep fragmentation and alterations expressed both in the EOG and EEG. Also, the findings of this project are in line with the hypothesis that iRBD is an intermediate stage between controls and PD patients, and that analysis of sleep signals hold huge potential for identification of PD biomarkers. The project illustrates the challenges associated with identifying a reliable biomarker that is both sensitive and specific. Finally, the project highlights how appropriate biomedical signal processing can be used to reveal biomarkers in a robust and standardized way, and thereby support the evaluation of sleep from iRBD and PD patients.

The path and pace of the neurodegenerative progress varies between patients, and as the affection on specific neurons and brain areas differs, the characteristics of the different biomarkers are likely to vary. One patient may show decreased density and altered sleep spindles but normal REM stability and shifts between sleep stages, while the opposite is true for another patient. Also, the biomarkers provide a snapshot of the circumstances as they are based on only one night of sleep. To overcome the high inter-subject variability, the patients should ideally be referenced to themselves, so the trends of the biomarkers are investigated instead of the snapshots.

It is hard to set any limit values for abnormality based on one recording of PSG, and in spite of the indicative findings of this project, it may be unrealistic to think that one electrophysiological biomarker can stand alone and identify people who will develop PD. By combining several biomarkers, the precision of early disease identification is expected to raise simply because more aspects of a complex disease would be reflected. Especially if the biomarkers express alterations in different areas or mechanisms of the brain. Optimally, a combination of biomarker trends could indicate at which stage the disease is and at which pace and in which direction it is progressing. This would not only provide guidance to identify the patients at highest risk of developing PD, and thereby maybe facilitate PD treatment before PD diagnosis, but also provide insight in treatment efficiency and thereby help in personalized treatment.

PD is a complex disease and the progress, severity and symptom profile vary greatly across patients. The disease is sporadic and the symptom profile dependents on which brain areas and to what degree the neurodegeneration strikes. It is not known whether the diversity across PD patients are due to the sporadic nature of the disease, or if there exist PD subgroups pathological different. Questions such as why some PD patients do not have RBD nor RSWA, why some iRBD patients do not progress into PD or another NDD and why the iRBD patient group is much more male-dominated compared to the PD patient group are just few questions that lack answers. Although this project focused on synucleinopathy, some of the patients evaluated may suffer from a different pathology just not defined nor clarified yet.

It should be emphasized that the findings of this dissertation are all found through guidance using hypotheses built on how the neurodegeneration in PD could change sleep patterns. This dissertation has thus confirmed the theories by scientific proves and by generalized methods rather than discovered new physiologic findings. Additionally, all the findings are based on group comparisons revealing overall tendencies rather than specific indications of disease. The findings need to be validated with much more data, both from patients with iRBD, PD and other NDDs, but also from patients with sleep disorders such as apnea, narcolepsy and insomnia. Conclusively, although some of the here presented classifiers obtained high sensitivity and specificity, none of the suggested biomarkers would be able to identify early stages of neurodegeneration in a new individual with high enough certainty to be used as a diagnostic test.

6.1 Future aspects

Analysis of PSG recordings hold huge potential of early identification of PD, but is still a relatively new research area and a lot of investigation is still needed. Dividing in two categories, further investigation should aim to:

Medical aims:

- obtain prognostic data to investigate subject-specific trends for the biomakers.
- include untreated patients with PD to investigate the dopaminergic treatment effect.
- include patients with other NDDs or sleep disorders to test the usability of the biomarkers in a clinical setting.
- include multiscored data to compare inter-scorer reliability across pathologies and thereby make a better "gold standard" to train detectors from.
- include data from other sleep clinics in order to investigate the usability of the biomarkers across cites and countries.
- include more in-depth cognitive tests and disease severity measures for the patients as well as for the controls.

Technical aims:

- combine the potential biomarkers to investigate their joint strength for identifying patients with iRBD and PD.
- validate indicative EEG and EOG words given as input to the topic models in order to better understand what they reflect and their impact on different stages.
- investigate the continuous transitions between topics more thoroughly.
- investigate the sleep topic mixtures more thoroughly to investigate if patients show abnormal topic mixtures compared to controls.
- build new topic models where other words are tried, e.g. words describing SS, K-complexes, EMG tonus, rapid and slow EMs, etc.
- build sleep models released from the 30-s epoch rule in order to see the pace of sleep stage switching determined by a data-driven approach.
- build sleep models released from the five sleep stage rule in order to see how many sleep stages a data-driven approach suggests.
- build subject-specific or group-specific sleep models to understand and investigate the huge variability in sleep signals.

Bibliography

[1] A. Abbott. Neuroscience: while you were sleeping. *Nature*, 437(7063):1220–1222, 2005.

[2] American Academy of Sleep Medicine. *International classification of sleep disorders - third edition (ICSD-3)*. American Academy of Sleep Medicine, Westchester, IL, 3rd edition, 2014.

[3] D. Barber. *Bayesian reasoning and machine learning*. Cambridge University Press, New York, 2012.

[4] P. M. Bentley and J. T. E. McDonnell. Wavelet transforms : an introduction. *Electronics & Communication Engineering Journal*, 6(4):175–186, 1994.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[6] H. Braak, K. D. Tredici, U. Rüb, R. A. de Vos, E. N. H. J. Steur, and E. Braak. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*, 24(2):197–211, 2003.

[7] R. E. Brown, R. Basheer, J. T. McKenna, R. E. Strecker, and R. W. McCarley. Control of sleep and wakefulness. *Physiological Reviews*, 92(3):1087–1187, 2012.

[8] R. H. S. Carpenter. The neural control of looking. *Current Biology*, 10(8):291–293, 2000.

[9] S. S. Cash, E. Halgren, N. Dehghani, A. O. Rossetti, T. Thesen, C. Wang, O. Devinsky, R. Kuzniecky, W. Doyle, J. R. Madsen, E. Bromfield, L. Eröss, P. Halász, G. Karmos, R. Csercsa, L. Wittner, and I. Ulbert. The human K-complex represents an isolated cortical down-state. *Science*, 324(5930):1084–1087, 2009.

[10] M. H. Chase, P. J. Soja, and F. R. Morales. Evidence that glycine mediates the postsynaptic potentials that inhibit lumbar motoneurons during the atonia of active sleep. *The Journal of Neuroscience*, 9(3):743–751, 1989.

[11] K. R. Chaudhuri, D. G. Healy, and A. H. Schapira. Non-motor symptoms of Parkinson's disease: diagnosis and management. *The Lancet Neurology*, 5(3):235–245, 2006.

[12] C. Chen, Z.-M. Zhang, M.-L. Ouyang, X. Liu, L. Yi, Y.-Z. Liang, and C.-P. Zhang. Shrunken centroids regularized discriminant analysis as a promising strategy for metabolomics data exploration. *Journal of Chemometrics*, 29(3):154–164, 2015.

[13] H. F. Chiu and Y. K. Wing. REM sleep behaviour disorder: an overview. *International Journal of Clinical Practice*, 51(7):451–454, 1997.

[14] J. A. E. Christensen, R. Frandsen, J. Kempfner, L. Arvastson, S. R. Christensen, P. J. Jennum, and H. B. D. Sorensen. Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel. In *34th Annual International Conference of the IEEE EMBS*, pages 2941–2944, 2012.

[15] J. A. E. Christensen, J. Kempfner, M. Zoetmulder, H. L. Leonthin, L. Arvastson, S. R. Christensen, H. B. D. Sorensen, and P. Jennum. Decreased sleep spindle density in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. *Clinical Neurophysiology*, 125(3):512–519, 2014.

[16] J. A. E. Christensen, H. Koch, R. Frandsen, J. Kempfner, L. Arvastson, S. R. Christensen, H. B. D. Sorensen, and P. Jennum. Classification of iRBD and Parkinson's disease patients based on eye movements during sleep. In *35th Annual International Conference of the IEEE EMBC*, pages 441–444, 2013.

[17] I. M. Colrain and K. B. Campbell. The use of evoked potentials in sleep research. *Sleep Medicine Reviews*, 11(4):277–293, 2007.

[18] N. Cristianini and J. Shawe-Taylor. *An Introduction of Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 1st edition, 2000.

[19] K. Crowley. Sleep and sleep disorders in older adults. *Neuropsychology Review*, 21(1):41–53, 2011.

[20] K. Crowley, J. Trinder, Y. Kim, M. Carrington, and I. M. Colrain. The effects of normal aging on sleep spindle and K-complex production. *Clinical Neurophysiology*, 113(10):1615–1622, 2002.

[21] H. Danker-Hopfe, D. Kunz, G. Gruber, G. Klösch, J. L. Lorenzo, S. L. Himanen, B. Kemp, T. Penzel, J. Röschke, H. Dorn, A. Schlögl, E. Trenker, and G. Dorffner. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *Journal of Sleep Research*, 13(1):63–69, 2004.

[22] L. De Gennaro, M. Ferrara, and M. Bertini. The spontaneous K-complex during stage 2 sleep: Is it the 'forerunner' of delta waves? *Neuroscience Letters*, 291(1):41–43, 2000.

[23] L. M. L. de Lau and M. M. B. Breteler. Epidemiology of Parkinson's disease. *The Lancet Neurology*, 5(6):525–535, 2006.

[24] A. B. Dos Santos, K. A. Kohlmeier, and G. E. Barreto. Are Sleep Disturbances Preclinical Markers of Parkinson's Disease? *Neurochemical Research*, 40(3):421–427, 2015.

[25] M. L. Fantini, J.-F. Gagnon, D. Petit, S. Rompré, A. Décary, J. Carrier, and J. Montplaisir. Slowing of electroencephalogram in rapid eye movement sleep behavior disorder. *Annals of Neurology*, 53(6):774–780, 2003.

[26] J. A. Floyd, S. M. Medler, J. W. Ager, and J. J. Janisse. Age-related changes in initiation and maintenance of sleep: a meta-analysis. *Research in Nursing & Health*, 23(2):106–117, 2000.

[27] S. M. Fogel, N. Martin, M. Lafortune, M. Barakat, K. Debas, S. Laventure, V. Latreille, J.-F. Gagnon, J. Doyon, and J. Carrier. NREM sleep oscillations and brain plasticity in aging. *Frontiers in Neurology*, 3(176):1–7, 2012.

[28] S. M. Fogel and C. T. Smith. The function of the sleep spindle: a physiological index of intelligence and a mechanism for sleep-dependent memory consolidation. *Neuroscience and Biobehavioral Reviews*, 35(5):1154–1165, 2011.

[29] J. E. Galvin, M.-Y. Lee, and J. Q. Trojanowski. Synucleinopathies: clinical and pathological implications. *Archives of Neurology*, 58(2):186–190, 2001.

[30] L. D. Gennaro and M. Ferrara. Sleep spindles : an overview. *Sleep Medicine*, 7(5):423–440, 2003.

[31] N. Guazzelli, I. Feinberg, M. Aminoff, G. Fein, T. Floyd, and C. Maggini. Sleep spindles in normal elderly: Comparison with young adult patterns and relation to nocturnal awakening, cognitive function and brain atrophy. *Electroencephalography and Clinical Neurophysiology*, 63:526–539, 1986.

[32] Y. Guo, T. Hastie, and R. Tibshirani. Regularized linear discriminant analysis and its application in microarrays. *Biostatistics*, 8(1):86–100, 2007.

[33] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer, New York, 2nd edition, 2008.

[34] C. Iber, S. Ancoli-Israel, A. L. Chesson, and S. F. Quan. *The AASM Manual for the Scoring of Sleep and Associated Events: rules, terminology, and technical specification.* American Academy of Sleep Medicine, Westchester, IL, 2007.

[35] Y. Inoue, T. Sasai, and K. Hirata. Electroencephalographic finding in idiopathic REM sleep behavior disorder. *Neuropsychobiology*, 71(1):25–33, 2015.

[36] A. Iranzo. Sleep-wake changes in the premotor stage of Parkinson disease. *Journal of the Neurological Sciences*, 310(1-2):283–285, 2011.

[37] A. Iranzo, V. Isetta, J. L. Molinuevo, M. Serradell, D. Navajas, R. Farre, and J. Santamaria. Electroencephalographic slowing heralds mild cognitive impairment in idiopathic REM sleep behavior disorder. *Sleep Medicine*, 11(6):534–539, 2010.

[38] A. Iranzo, J. L. Molinuevo, J. Santamaría, M. Serradell, M. J. Martí, F. Valldeoriola, and E. Tolosa. Rapid-eye-movement sleep behaviour disorder as an early marker for a neurodegenerative disorder: a descriptive study. *The Lancet Neurology*, 5(7):572–577, 2006.

[39] P. S. Jensen, H. B. D. Sorensen, H. L. Leonthin, and P. Jennum. Automatic sleep scoring in normals and in individuals with neurodegenerative disorders according to new international sleep scoring criteria. *Journal of Clinical Neurophysiology*, 27(4):296–302, 2010.

[40] S.-H. Kang, I.-Y. Yoon, S. D. Lee, J. W. Han, T. H. Kim, and K. W. Kim. REM sleep behavior disorder in the Korean elderly population: prevalence and clinical characteristics. *Sleep*, 36(8):1147–1152, 2013.

[41] J. Kempfner, G. L. Sorensen, M. Nikolic, R. Frandsen, H. B. D. Sorensen, and P. Jennum. Rapid eye movement sleep behavior disorder as an outlier detection problem. *Journal of Clinical Neurophysiology*, 31(1):86–93, 2014.

[42] H. Koch, J. A. E. Christensen, R. Frandsen, M. Zoetmulder, S. R. Christensen, L. Arvastson, P. Jennum, and H. B. D. Sorensen. Automatic sleep classification using a data-driven topic model reveals latent sleep states. *Journal of Neuroscience Methods*, 235:130–137, 2014.

[43] M. H. Kryger, T. Roth, and W. C. Dement. *Principles and Practice of Sleep Medicine*. Elseiver Inc., 5th edition, 2010.

[44] P. Y. Ktonas, S. Golemati, P. Xanthopoulos, V. Sakkalis, M. D. Ortigueira, H. Tsekou, M. Zervakis, T. Paparrigopoulos, A. Bonakis, N. T. Economou, P. Theodoropoulos, S. G. Papageorgiou, D. Vassilopoulos, and C. R. Soldatos. Time-frequency analysis methods to quantify the time-varying microstructure of sleep EEG spindles: possibility for dementia biomarkers? *Journal of Neuroscience Methods*, 185(1):133–142, 2009.

[45] J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.

[46] V. Latreille, J. Carrier, M. Lafortune, R. B. Postuma, J.-A. Bertrand, M. Panisset, S. Chouinard, and J.-F. Gagnon. Sleep spindles in Parkinson's disease may predict the development of dementia. *Neurobiology of Aging*, 36(2):1083–1090, 2015.

[47] V. Latreille, J. Carrier, J. Montplaisir, M. Lafortune, and J. F. Gagnon. Non-rapid eye movement sleep characteristics in idiopathic REM sleep behavior disorder. *Journal of the Neurological Sciences*, 310(1-2):159–162, 2011.

[48] S. C. Leiser, J. Dunlop, M. R. Bowlby, and D. M. Devilbiss. Aligning strategies for using EEG as a surrogate biomarker: a review of preclinical and clinical research. *Biochemical Pharmacology*, 81(12):1408–1421, 2011.

[49] P. Lindgren, S. V. Campenhausen, E. Spottke, U. Siebert, and R. Dodel. Cost of Parkinson's disease in Europe. *European Journal of Neurology*, 12(Suppl 1):68–73, 2005.

[50] P.-H. Luppi, O. Clément, and P. Fort. Paradoxical (REM) sleep genesis by the brainstem is under hypothalamic control. *Current Opinion in Neurobiology*, 23(5):786–792, 2013.

[51] P.-H. Luppi, O. Clément, E. Sapin, D. Gervasoni, C. Peyron, L. Léger, D. Salvert, and P. Fort. The neuronal network responsible for paradoxical sleep and its dysfunctions causing narcolepsy and rapid eye movement (REM) behavior disorder. *Sleep Medicine Reviews*, 15(3):153–163, 2011.

[52] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993.

[53] J. Massicotte-Marquez, J. Carrier, A. Décary, A. Mathieu, M. Vendette, D. Petit, and J. Montplaisir. Slow-wave sleep and delta power in rapid eye movement sleep behavior disorder. *Annals of Neurology*, 57(2):277–282, 2005.

[54] C. Micanovic and S. Pal. The diagnostic utility of EEG in early-onset dementia: a systematic review of the literature with narrative analysis. *Journal of Neural Transmission*, 121(1):59–69, 2014.

[55] S. L. Naismith, S. J. G. Lewis, and N. L. Rogers. Sleep-wake changes and cognition in neurodegenerative disease. *Progress in Brain Research*, 190:21–52, 2011.

[56] L. Nguyen, J. L. Bradshaw, J. C. Stout, R. J. Croft, and N. Georgiou-Karistianis. Electrophysiological measures as potential biomarkers in Huntington's disease: review and future directions. *Brain Research Reviews*, 64(1):177–194, 2010.

[57] C. L. Nicholas, J. Trinder, K. E. Crowley, and I. M. Colrain. The impact of slow wave sleep proximity on evoked K-complex generation. *Neuroscience Letters*, 404(1-2):127–131, 2006.

[58] A. Nicolas, D. Petit, S. Rompré, and J. Montplaisir. Sleep spindle characteristics in healthy subjects of different age groups. *Clinical Neurophysiology*, 112(3):521–527, 2001.

[59] R. L. Nussbaum and C. E. Ellis. Alzheimer's disease and Parkinson's disease. *The New England Journal of Medicine*, 14(348):1356–1364, 2003.

[60] M. M. Ohayon, M. A. Carskadon, C. Guilleminault, and M. V. Vitiello. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep*, 27(7):1255–1273, 2004.

[61] M. M. Ohayon, M. Caulet, and R. Priest. Violent behavior during sleep. *Journal of Clinical Psychiatry*, 58(8):369–376, 1997.

[62] R. B. Postuma, J. F. Gagnon, S. Rompré, and J. Y. Montplaisir. Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease. *Neurology*, 74(3):239–244, 2010.

[63] G. Rauchs, M. Schabus, S. Parapatics, F. Bertran, P. Clochon, P. Hot, P. Denise, B. Desgranges, F. Eustache, G. Gruber, and P. Anderer. Is there a link between sleep changes and memory in Alzheimer's disease? *Neuroreport*, 19(11):1159–1162, 2008.

[64] J. Rodrigues Brazète, J. Montplaisir, D. Petit, R. B. Postuma, J. A. Bertrand, D. Génier Marchand, and J. F. Gagnon. Electroencephalogram slowing in rapid eye movement sleep behavior disorder is associated with mild cognitive impairment. *Sleep Medicine*, 14(11):1059–1063, 2013.

[65] K. Sakai, S. Crochet, and H. Onoe. Pontine structures and mechanisms involved in the generation of paradoxical (REM) sleep. *Archives Italiennes de Biologie*, 139(1-2):93–107, 2001.

[66] C. B. Saper, T. C. Chou, and T. E. Scammell. The sleep switch: hypothalamic control of sleep and wakefulness. *Trends in Neurosciences*, 24(12):726–731, 2001.

[67] C. B. Saper, P. M. Fuller, N. P. Pedersen, J. Lu, and T. E. Scammell. Sleep State Switching. *Neuron*, 68(6):1023–1042, 2010.

[68] T. Sasai, M. Matsuura, and Y. Inoue. Electroencephalographic findings related with mild cognitive impairment in idiopathic rapid eye movement sleep behavior disorder. *Sleep*, 36(12):1893–1899, 2013.

[69] M. Schabus, K. Hödlmoser, G. Gruber, C. Sauter, P. Anderer, G. Klösch, S. Parapatics, B. Saletu, W. Klimesch, and J. Zeitlhofer. Sleep spindle-related activity in the human EEG and its relation to general cognitive and learning abilities. *The European Journal of Neuroscience*, 23(7):1738–1746, 2006.

[70] C. H. Schenck, B. F. Boeve, and M. W. Mahowald. Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye

movement sleep behavior disorder: a 16-year update on a previously reported series. *Sleep Medicine*, 14(8):744–748, 2013.

[71] C. H. Schenck, S. R. Bundlie, and M. W. Mahowald. Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behaviour disorder. *Neurology*, 46(2):388–393, 1996.

[72] C. H. Schenck, S. R. Bundlie, and M. W. Mahowald. REM behavior disorder (RBD): delayed emergence of parkinsonism and/or dementia in 65% of older men initially diagnosed with idiopathic RBD, and an analysis of the minimum & maximum tonic and/or phasic electromyographic abnormalities found during REM sleep. *Sleep*, 26:A316–A316, 2003.

[73] C. H. Schenck, J. Y. Montplaisir, B. Frauscher, B. Hogl, J.-F. Gagnon, R. Postuma, K. Sonka, P. Jennum, M. Partinen, I. Arnulf, V. Cochen de Cock, Y. Dauvilliers, P.-H. Luppi, A. Heidbreder, G. Mayer, F. Sixel-Döring, C. Trenkwalder, M. Unger, P. Young, Y. K. Wing, L. Ferini-Strambi, R. Ferri, G. Plazzi, M. Zucconi, Y. Inoue, A. Iranzo, J. Santamaria, C. Bassetti, J. C. Möller, B. F. Boeve, Y. Y. Lai, M. Pavlova, C. Saper, P. Schmidt, J. M. Siegel, C. Singer, E. St Louis, A. Videnovic, and W. Oertel. Rapid eye movement sleep behavior disorder: devising controlled active treatment studies for symptomatic and neuroprotective therapy–a consensus statement from the International Rapid Eye Movement Sleep Behavior Disorder Study Group. *Sleep Medicine*, 14(8):795–806, 2013.

[74] J. R. L. Schwartz and T. Roth. Neurophysiology of sleep and wakefulness: basic science and clinical implications. *Current Neuropharmacology*, 6(4):367–378, 2008.

[75] G. L. Sorensen, J. Kempfner, M. Zoetmulder, H. B. D. Sorensen, and P. Jennum. Attenuated heart rate response in REM sleep behavior disorder and Parkinson's disease. *Movement Disorders*, 27(7):888–894, 2012.

[76] G. L. Sorensen, J. Mehlsen, and P. Jennum. Reduced sympathetic activity in idiopathic rapid-eye-movement sleep behavior disorder and Parkinson's disease. *Autonomic Neuroscience: Basic and Clinical*, 179(1-2):138–141, 2013.

[77] T. L. Sorensen, U. L. Olsen, I. Conradsen, J. Henriksen, T. W. Kjaer, C. E. Thomsen, and H. B. D. Sorensen. Automatic epileptic seizure onset detection using Matching pursuit: a case study. In *32nd Annual International Conference of the IEEE EMBS*, pages 3277–3280, 2010.

[78] M. Steriade, D. McCormick, and T. Sejnowski. Thalamocortical oscillations in the sleeping and aroused brain. *Science*, 262(5134):679–685, 1993.

[79] M. Steriade and I. Timofeev. Neuronal plasticity in thalamocortical networks during sleep and waking oscillations. *Neuron*, 37(4):563–576, 2003.

[80] K. Stiasny-Kolster, G. Mayer, S. Schäfer, J. C. Möller, M. Heinzel-Gutenbrunner, and W. H. Oertel. The REM sleep behavior disorder screening questionnaire-a new diagnostic instrument. *Movement Disorders*, 22(16):2386–2393, 2007.

[81] R. Tibshirani. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodological)*, 58(1):267–288, 1996.

[82] A. C. Vernon, C. Ballard, and M. Modo. Neuroimaging for Lewy body disease: Is the in vivo molecular imaging of α-synuclein neuropathology required and feasible? *Brain Research Reviews*, 65(1):28–55, 2010.

[83] J. Wang, J. G. Hoekstra, C. Zuo, T. J. Cook, and J. Zhang. Biomarkers of Parkinson's disease: current status and future. *Drug Discovery Today*, 18(3-4):155–162, 2013.

[84] S. C. Warby, S. L. Wendt, P. Welinder, E. G. S. Munk, O. Carrillo, H. B. D. Sorensen, P. Jennum, P. E. Peppard, P. Perona, and E. Mignot. Sleep-spindle detection: crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nature Methods*, 11(4):385–392, 2014.

[85] C. E. Westerberg, B. A. Mander, S. M. Florczak, S. Weintraub, M.-M. Mesulam, P. C. Zee, and K. A. Paller. Concurrent impairments in sleep and memory in amnestic mild cognitive impairment. *Journal of the International Neuropsychological Society*, 18(3):490–500, 2012.

[86] S. Zoccolella, M. Savarese, P. Lamberti, R. Manni, C. Pacchetti, and G. Logroscino. Sleep disorders and the natural history of Parkinson's disease: The contribution of epidemiological studies. *Sleep Medicine Reviews*, 15(1):41–50, 2011.

[87] J. Zygierewicz, K. J. Blinowska, P. J. Durka, W. Szelenberger, S. Niemcewicz, and W. Androsiuk. High resolution study of sleep spindles. *Clinical Neurophysiology*, 110(12):2136–2147, 1999.

Appendices

Paper I

TITLE:
Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel

AUTHORS:
Julie A. E. Christensen, Rune Frandsen, Jacob Kempfner, Lars Arvastson, Søren R. Christensen, Poul Jennum and Helge B. D. Sorensen

JOURNAL:
34th Annual International Conference of the IEEE EMBC

YEAR:
2012

VOLUME:
2012

PAGES:
2941-4

PUBLICATION HISTORY:
Accepted as poster presentation at the 34th Annual International Conference of the IEEE Engineering in medicine and Biology Society (EMBC), San Diego, CA USA, 28 August - 1 September, 2012

34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel

Julie A.E. Christensen[a,b,c], Rune Frandsen[c], Jacob Kempfner[a,c], Lars Arvastson[b], Soren R. Christensen[b], Poul Jennum[c] and Helge B.D. Sorensen[a]

Abstract— **In this study, polysomnographic left side EOG signals from ten control subjects, ten iRBD patients and ten Parkinson's patients were decomposed in time and frequency using wavelet transformation. A total of 28 features were computed as the means and standard deviations in energy measures from different reconstructed detail subbands across all sleep epochs during a whole night of sleep. A subset of features was chosen based on a cross validated Shrunken Centroids Regularized Discriminant Analysis, where the controls were treated as one group and the patients as another. Classification of the subjects was done by a leave-one-out validation approach using same method, and reached a sensitivity of 95%, a specificity of 70% and an accuracy of 86.7%. It was found that in the optimal subset of features, two hold lower frequencies reflecting the rapid eye movements and two hold higher frequencies reflecting EMG activity. This study demonstrates that both analysis of eye movements during sleep as well as EMG activity measured at the EOG channel hold potential of being biomarkers for Parkinson's disease.**

I. INTRODUCTION

It has been stated that patients suffering from the sleep disorder idiopathic Rapid Eye Movement (REM) Sleep Behavior Disorder (iRBD) are at high risk of developing Parkinson's disease (PD) [1], which makes them essential to analyze in the search for biomarkers of PD. In consequence, many studies focus on sleep data in the search for biomarkers, where polysomnographic (PSG) data have been analysed, including analysis of electromyography (EMG) [2], electroencephalography (EEG) [3] and sleep variables such as sleep latency, sleep time, percentage distribution of sleep stages and sleep efficiency [2].

According to Braak et al., the evolution of PD will involve the basale brain structures to start with (Braak stage I-II), and thereafter progress to the additional brain regions (Braak stage III-IV) [4]. During sleep, eye movements (EMs) are controlled by neurons located in the brain stem structures. In this study, it is therefore hypothesized that patients with iRBD and especially patients with PD will reflect abnorm form of EMs during sleep. To the best knowledge of the authors, no other studies have been focusing on analyzing EMs measured as electrooculography (EOG) during sleep, and for that reason, the study presented here is a pilot study revealing whether EMs hold potential of being a biomarker for PD or not.

The recording of EMs is done by EOG, which is based on a potential difference between the anterior (cornea) and the posterior (retina) point of the eyeball. In that way, the eye acts as a dipole in which the cornea is positive and the retina is negative. By placing electrodes besides each outer canthus, the EMs will be registered as positive potentials by the electrode nearest the cornea and as negative potentials by the electrode nearest the retina. Because of the simultaneously movement of the eyeballs, the EMs registered at the left and right EOG electrode will always appear synchronic and anti-correlated. For clarity, EMs during sleep are in this study defined as holding slow and fast EMs (SEMs and REMs), where the main part of the SEMs lie in the range 0.5-1 Hz and the main part of the REMs lie in the range 1-5 Hz.

II. DATA ACQUISITION

A. Subjects

The patients enrolled in this study were evaluated at the Danish Center for Sleep Medicine at Glostrup Hospital in Denmark. The evaluation of the patients included PSG, multiple sleep latency test and a comprehensive medical history and medication. Patients taking any anti-depressant drug, including hypnotics were excluded, though dopaminergic treatment was continued. Also, the quality of the PSG data was individually evaluated. If too much noise, such as disconnection, was present on the recordings making either the sleep stage scoring or the further analysis unreliable, the subject was excluded. A total of ten PD patients and ten iRBD patients were included in this study. Furthermore, ten age-matched control subjects without history of movement disorder, dream enacting behavior or other former diagnosed sleep disorders were included as controls. Additionally, no medication known to affect sleep was acceptable.

B. Polysomnographic recordings

All controls underwent at least one night of PSG recorded outpatient, and all patients underwent at least one night of PSG recorded either outpatient or in-hospital. For the outpatient recordings, the PSG equipment was fitted at the clinic. The PSG recordings were performed in accordance with the sleep scoring standard stated in 2004 by the American Academy of Sleep Medicine (AASM) [5]. The EOG

Corresponding author: Julie AEC: jaec@elektro.dtu.dk
Contact information: HBDS: hbs@elektro.dtu.dk
RF: RFRA0011@regionh.dk, JK: jkem@elektro.dtu.dk
LA: larv@lundbeck.com, SRC: srac@lundbeck.com,
and PJ: poje@glo.regionh.dk
[a]DTU Electrical Eng., Ørsteds Plads, bldg. 349, DK-2800 Kgs. Lyngby
[b]H. Lundbeck A/S, Ottiliavej 9, DK-2500 Valby
[c]Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Center for Healthy Aging, Faculty of Health Sciences, University of Copenhagen, Glostrup Hospital, DK-2600 Glostrup

978-1-4577-1787-1/12/$26.00 ©2012 IEEE

electrodes were placed one cm out and up (left) or down (right) from the outer canthus with reference to the mastoids.

The sleep staging of all subjects were performed by experienced PSG technicians in accordance with the AASM standard [5] staging every epoch of 30 seconds of PSG data into either REM sleep, three stages of non-REM sleep (N1, N2 or N3) or wake (W) resulting in a hypnogram of same length as the entire recording. The total number of scored epochs between lights off and lights on is seen in Table I.

TABLE I

TOTAL NO. OF EPOCHS BETWEEN LIGHTS OFF AND LIGHTS ON.

Stage	Controls	iRBD	PD
Wake (%)	1173 (12)	1881 (18)	1882 (19)
REM (%)	2000 (21)	1731 (16)	1531 (15)
N1 (%)	678 (7)	1081 (10)	1275 (13)
N2 (%)	4443 (46)	4881 (46)	4073 (42)
N3 (%)	1347 (14)	1114 (10)	1084 (11)
Sum (Σ%)	9641 (100)	10688 (100)	9827 (100)

The raw sleep data, hypnograms and sleep events were extracted from Nervus (V5.7, Cephalon DK, Nørresundby, Denmark) using the build-in export data tool. For further analysis, the data were imported to MATLAB (R2012a, The MathWorks, Natick, MA, USA). The analyzed data had a sampling frequency of 256 Hz.

III. METHODOLOGY

A schematic illustration of the signal processing in this study is seen in Fig. 1. For feature extraction, the wavelet decomposition technique was used, and the Shrunken Centroids Regularized Discriminant Analysis (SCRDA) method was used to choose a subset of features and classify the subjects [6]. The wavelet decomposition technique was chosen as we preferred a good resolution at lower frequencies compared to higher frequencies, as the lower frequency ranges are the ones reflecting EMs during sleep. Other feature extraction techniques such as Principal Component Analysis (PCA) or Singular Value Decomposition (SVD) could have been tried, but as this is a preliminary study, we preferred to use a method in which the features are easily interpreted. The SCRDA was chosen as it is designed for classification problems where the number of features is larger or nearly the same as the number of samples. Additionaly, it is suitable for feature elimination purposes, which enabled us to choose the optimal subset of features. Below follows a more detailed description of the steps seen in Fig. 1.

A. Feature extraction

For feature extraction, only the left side EOG signal (EOGL) was used. Before the feature extraction, the 50 Hz powerline noise was removed by a FIR equiripple notch filter with cutoff frequencies at 49 Hz and 51 Hz, respectively. The hypnogram and the EOGL signal were evaluated in all sleep stages (both REM and non-REM) in between lights off and lights on. Each sleep epoch was decomposed using the

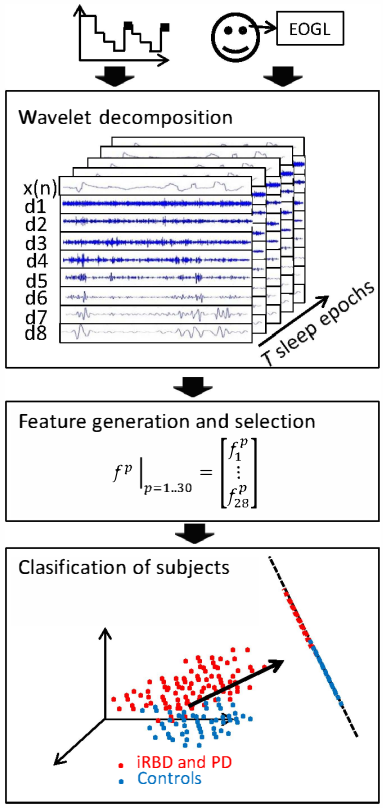


Fig. 1. A schematical view of the methodology of this study. Each sleep epoch between lights off and lights on was extracted from the left side EOG signal and decomposed using wavelet decomposition. The energy percentages and the common logarithm of the summed absolute signal values of the reconstructed detail subbands d2-d8 were computed. The mean and standard deviation of these across all sleep epochs composed the feature vectors. An optimal subset of features were found and the subjects were classfied by use of the Shrunken Centroids Regularized Discriminant Analysis (SCRDA) method.

Discrete Wavelet Transform (DWT). In DWT, an input signal $x(n)$ is passed through a series of filters, which split the signal equally into its high- and low frequency components denoted as detail (D) and approximation (A) components, respectively [7]. After filtering, the components contain redundancies and therefore a downsampling by a factor of two is applied. The decomposition of each sleep epoch was carried out to level eight using a Daubechies 4 (db4) as the mother wavelet. The mother wavelet can be described as the structure which is used to analyze the input signal by comparing it with different scaled and shifted mother wavelets [7]. The choice of the db4 was based on other studies analyzing EOG, [8] [9], as well as a wish to have a rather smooth basisfunction.

Following the wavelet decomposition, the energy percentages and the common logarithm of the summed absolute signal values of the reconstructed detail subbands d2-d8 were computed, yielding 14 values for each sleep epoch.

A single feature vector for each subject was computed by taking the mean and standard deviation of the values across every sleep epoch, thereby yielding f^p holding 28 feature values describing subject p

$$f^p = \begin{bmatrix} \left[f^p_{\mu(\%E)}\right]^{d8}_{d2} \\ \left[f^p_{\sigma(\%E)}\right]^{d8}_{d2} \\ \left[f^p_{\mu(\log_{10}E)}\right]^{d8}_{d2} \\ \left[f^p_{\sigma(\log_{10}E)}\right]^{d8}_{d2} \end{bmatrix} \quad \text{where} \quad \left[f^p_E\right]^{d8}_{d2} = \begin{bmatrix} f_{E_{d2}} \\ \vdots \\ f_{E_{d8}} \end{bmatrix}, \quad (1)$$

where $\%E_{dx}$ and $\log_{10}E_{dx}$ indicate the energy percentage and the common logarithm of the summed absolute signal values of the reconstructed detail subband dx, respectively. The σ and μ indicate the mean and standard deviation across all sleep epochs, respectively. These measures were computed in trying to reflect the hypothesis that the patients have low amplitude EMs during long periods of the night compared to the controls who have concentrated periods with pronaunced EMs.

B. Feature evaluation and selection

The 30 subjects included in this study were each represented by a feature vector of 28 feature values. To avoid overfitting, feature evaluation was therefore needed to appropriately choose a subset of features. This was done by use of the SCRDA method, which generalizes the idea of the Nearest Shrunken Centroids (NSC) into the classical discriminant analysis [6]. In the standard Linear Discriminant Analysis (LDA), G populations are assumed to have a multivariate normal distribution and they are described by mean vectors μ_g $(g = 1, ..., G)$ and a common covariance matrix Σ, [6]. An observation $x_{g,i}$ is classified to a population g which minimizes $(x_{g,i} - \mu_g)^T \Sigma^{-1}(x_{g,i} - \mu_g)$, which under the multivariate normal assumptions is the same as chosing the population that maximizes the likelihood of the observation. When including the prior probabilities π_g of each population, the choice of population can be stated based on the posterior probability rather than the likelihood, and because of the assumption of common covariance matrix, the criteria in LDA can be summarized to,

$$\begin{aligned} x_{g,i} \in \text{population}\left(g = \text{argmax}_{g'} d_{g'}(x_{g,i})\right) \quad \text{where} \\ d_g(x) = x^T \Sigma^{-1} \mu_g - \frac{1}{2} \mu_g^T \Sigma^{-1} \mu_g + \log(\pi_g). \end{aligned} \quad (2)$$

The discriminant function, $d_g(x)$, is used in a sample version, $\hat{d}_g(x) = x^T \hat{\Sigma}^{-1} \hat{\mu}_g - \frac{1}{2} \hat{\mu}_g^T \hat{\Sigma}^{-1} \hat{\mu}_g + \log \pi_g$ with $\hat{\mu}_g = \frac{1}{n_g} \sum_{i=1}^{n_g} x_{g,i}$ and $\hat{\Sigma} = \frac{1}{n} \left(X - \overline{X}\right) \left(X - \overline{X}\right)^T$. Here, X is a $p \times n$ matrix holding the observations columnwise, $\overline{X}$ is a $p \times n$ matrix holding the observation mean vectors columnwise and n_g is the number of samples in population g. The sample covariance matrix calculated in this way is badly estimated (or even singular) in the cases where the number of samples is small compared to the number of features. To overcome this issue, different forms of regularizations can be done in the estimation of the covariance matrix, and in SCRDA, the covariance matrix is regularized by use of parameter α, $0 \leq \alpha \leq 1$ as in equation 3,

$$\tilde{\Sigma} = \alpha \hat{\Sigma} + (1 - \alpha) \hat{D} \;\text{ with }\; \hat{D} = \text{diag}(\hat{\Sigma}) \quad (3)$$

In SCRDA, a threshold forces the group centroids of a particular feature towards zero [6]. In this way, features will be eliminated, and an optimal subset of features can be found. The idea is incorporated by substituting the centroids $\hat{\mu}_g$ by the shrunken centroids $\tilde{\mu}_g$ found by,

$$\tilde{\mu} = \text{sgn}(\tilde{\Sigma}^{-1} \hat{\mu}) \left(|\tilde{\Sigma}^{-1} \hat{\mu}| - \Delta\right)_+ . \quad (4)$$

Conclusively, the criteria in the SCRDA method both stabilizes the covariance matrix and eliminates non-contributing features and can be stated as,

$$\begin{aligned} x_{g,i} \in \text{population}\left(g = \text{argmax}_{g'} \tilde{d}_{g'}(x_{g,i})\right) \quad \text{where} \\ \tilde{d}_g(x) = x^T \tilde{\Sigma}^{-1} \tilde{\mu}_g - \frac{1}{2} \tilde{\mu}_g^T \tilde{\Sigma}^{-1} \tilde{\mu}_g + \log(\pi_g). \end{aligned} \quad (5)$$

Each subject was classified by a leave-one-out approach, where the optimal value for the threshold Δ and the regularizing parameter α was found by a 10-fold crossvalidation on the training set consisting of 29 subjects. The final values of Δ and α were found as the mean across the 30 runs. The mean values for Δ and α indirectly give the optimal subset of features across the 30 runs.

IV. RESULTS AND DISCUSSION

In Table II is seen the optimal subset of features found by the mean values for Δ and α across the 30 runs.

TABLE II

THE OPTIMAL SUBSET OF FEATURES FOUND BY THE SCRDA METHOD, WHERE THE PARAMETERS Δ AND α WERE FOUND AS THE MEAN ACROSS THE 30 RUNS IN THE LEAVE-ONE-OUT APPROACH. WITHIN EACH RUN, THE PARAMETERS WERE FOUND BY A 10-FOLD CROSS VALIDATION.

Original feature no.	Frequency range [Hz]	Description of feature
1	32-64	Mean of the logarithmic "energy" in d2
2	16-32	Mean of the logarithmic "energy" in d3
20	1-2	Mean of the percentage energy in d7
27	1-2	Std of the percentage energy in d7

It is seen that the optimal subset of features includes one feature derived from the reconstructed detail subband d2, one from d3 and two from d7. These reconstructed detail subbands hold frequencies in the range 32-64 Hz, 16-32 Hz and 1-2 Hz, respectively. Using the REM frequency range definition of 1-5 Hz as stated in [10], two of the four included features found in this study reflect REMs. According to the AASM standard, EOG signals should be evaluated in the frequency range 0.3-35 Hz and EMG signals in the frequency range 10-100 Hz when scoring sleep data [5]. Based on this, it can be stated that EOG activity free of EMG activity includes frequencies below 10 Hz. Feature number 2 could therefore hold a portion of EM activity, but could also hold

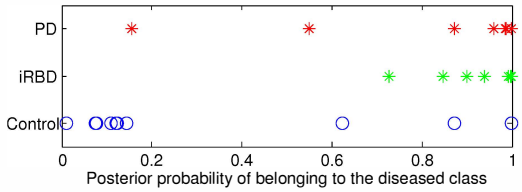


Fig. 2. The posterior probabilities of belonging to the diseased class. These were calculated for each subject during the leave-one-out classification. The diseased class holds the stars indicating iRBD (green) and PD (red) patients, and the control class holds the circles indicating control subjects. It is seen that by choosing the class with the highest posterior probability, three control subjects are misclassified as diseased and one PD patient is misclassified as control.

EMG activity, which must be considered more likely because of the higher frequency content. Feature number 1, on the other hand, must be concluded to hold much more EMG activity than EOG activity.

In the AASM standard, EEG signals should be evaluated in the frequency range 0.3-35 Hz, which is the same as the EOG signals [5]. In generel, EEG signals have less prominant amplitudes than EOG, and do therefore not appear in the EOG signals. During sleep, although, some EEG events such as distinct K-complexes and sleep spindles, are likely to appear in the EOG signals. During N3 sleep, EEG appears as slow wave activity (SWA), which is waves of frequencies 0.5-2 Hz and peak-to-peak amplitudes of above 75 μV [5]. SWA is mainly measured over the frontal regions, and is thereby very likely to appear in the EOG channels. It is not thoroughly evaluated how much of such EEG artifacts appears in the EOGL signal analyzed in this study. It is thereby neither evaluated in which degree such artifacts would affact the results obtained. It is though presumed, that the SWA is the main EEG artifact, that could have an impact on the results, and it is seen in Table I, that the amount of N3 sleep included in this study is more or less the same for each group.

Assuming that the EOG signals are anti-correlated during EMs, the presence of EEG artifacts may be reduced by analyzing a correlation measure between the two EOG signals. It should also be emphasized that the influence of other artifacts, such as baseline drift, has to be adressed in future work.

In Fig. 2 is seen the posterior probability of belonging to the diseased class. The posterior probabilities were calculated for each subject during the leave-one-out classification by use of the discriminant function in equation 5. The blue circles indicate the ten control subjects, the green stars indicate the ten iRBD patients and the red stars indicate the ten PD patients. When interpreting the result, it should be kept in mind that the iRBD and PD patients were treated as one class, i.e. the subset of features were found based on separation of controls and patients and not based on separation of controls, iRBD and PD patients.

Following the criteria in the SCRDA method stated in equation 5, it is seen from Fig. 2 that three control subjects are misclassified as diseased and one PD patient is misclassified as control. The classification approach in this study therefore obtain a sensitivity of 95%, a specificity of 70% and an accuracy of 86.7%.

It should be emphasized, that the four features presented in Table II was found by the mean values of the parameters Δ and α across the 30 runs in the leave-one-out classification. This means, that the included features for classifying each subject may vary as a consequence of the different parameter values in each run. Therefore, the classification results represented in Fig. 2 are not all obtained by use of the four features presented in Table II, as these reflect the overall best subset. During the simulations, though, it was noticed that feature 1 and 27 were represented in each of the 30 runs, indicating that these two features might be the most discriminative ones. This should, however, be evaluated more thoroughly in future work. Also, the patient group should be separated into iRBD and PD patients and it should be investigated which features would be selected in the three class case.

V. CONCLUSIONS

Selection of features holding different frequency bands yielded that the optimal subset of features include two features reflecting REMs and two features reflecting EMG activity. It is thus concluded, that EMs during sleep as well as EMG activity measured at one EOG channel hold information, which can be used to classify iRBD and PD patients. Allthough more research is needed, this study demonstrates that analysis of EMs during sleep and EMG activity during sleep both hold potential of being biomarkers for PD.

REFERENCES

[1] C. Schenck, S. Bundlie, and M. Mahowald, "Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behaviour disorder," *Neurology*, vol. 46, no. 2, pp. 388–393, 1996.

[2] R. B. Postuma, J. F. Gagnon, S. Rompré, and J. Y. Montplaisir, "Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease.," *Neurology*, vol. 74, pp. 239–244, Jan. 2010.

[3] V. Latreille, J. Carrier, J. Montplaisir, M. Lafortune, and J.-F. Gagnon, "Non-rapid eye movement sleep characteristics in idiopathic REM sleep behavior disorder.," *Journal of the Neurological Sciences*, vol. 310, pp. 159–162, June 2011.

[4] H. Braak, K. D. Tredici, U. Rüb, R. A. de Vos, E. N. H. J. Steur, and E. Braak, "Staging of brain pathology related to sporadic Parkinson's disease," *Neurobiology of Aging*, vol. 24, pp. 197–211, Apr. 2003.

[5] C. Iber, *The AASM Manual for the Scoring of Sleep and Associated Events*. American Academy of Sleep Medicine, 2007.

[6] Y. Guo, T. Hastie, and R. Tibshirani, "Regularized linear discriminant analysis and its application in microarrays," *Biostatistics*, vol. 8, no. 1, pp. 86–100, 2007.

[7] P. M. Bentley and J. T. E. McDonnell, "Wavelet transforms : an introduction," *Electronics & Communication Engineering Journal*, vol. 6, no. 4, pp. 175–186, 1994.

[8] J.-X. Ma, L.-C. Shi, and B.-L. Lu, "Vigilance estimation by using electrooculographic features.," in *32nd Annual International Conference of the IEEE EMBS*, pp. 6591–6594, Jan. 2010.

[9] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, "Eye Movement Analysis for Activity Recognition Using Electrooculography," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 741–753, 2011.

[10] R. Agarwal, T. Takeuchi, S. Laroche, and J. Gotman, "Detection of rapid-eye movements in sleep studies.," *IEEE Transactions on Bio-Medical Engineering*, vol. 52, pp. 1390–1396, Aug. 2005.

Paper II

TITLE:
Classification of iRBD and Parkinson's disease patients based on eye movements during sleep

AUTHORS:
Julie A. E. Christensen, Henriette Koch, Rune Frandsen, Jacob Kempfner, Lars Arvastson, Søren R. Christensen, Helge B. D. Sorensen and Poul Jennum

JOURNAL:
35th Annual International Conference of the IEEE EMBC

YEAR:
2013

VOLUME:
2013

PAGES:
441-4

PUBLICATION HISTORY:
Accepted as poster presentation at the 35th Annual International Conference of the IEEE Engineering in medicine and Biology Society (EMBC), Osaka, Japan, 3 July - 7 July, 2013

35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Classification of iRBD and Parkinson's Disease patients based on eye movements during sleep

Julie A. E. Christensen[a,b,c], Henriette Koch[a], Rune Frandsen[c], Jacob Kempfner[a,c], Lars Arvastson[b], Soren R. Christensen[b], Helge B. D. Sorensen[a] and Poul Jennum[c]

Abstract— **Patients suffering from the sleep disorder idiopathic rapid-eye-movement sleep behavior disorder (iRBD) have been observed to be in high risk of developing Parkinson's disease (PD). This makes it essential to analyze them in the search for PD biomarkers. This study aims at classifying patients suffering from iRBD or PD based on features reflecting eye movements (EMs) during sleep. A Latent Dirichlet Allocation (LDA) topic model was developed based on features extracted from two electrooculographic (EOG) signals measured as parts in full night polysomnographic (PSG) recordings from ten control subjects. The trained model was tested on ten other control subjects, ten iRBD patients and ten PD patients, obtaining a EM topic mixture diagram for each subject in the test dataset. Three features were extracted from the topic mixture diagrams, reflecting "certainty","fragmentation" and "stability" in the timely distribution of the EM topics. Using a Naive Bayes (NB) classifier and the features "certainty" and "stability" yielded the best classification result and the subjects were classified with a sensitivity of 95 %, a specificity of 80 % and an accuracy of 90 %. This study demonstrates in a data-driven approach, that iRBD and PD patients may exhibit abnorm form and/or timely distribution of EMs during sleep.**

I. INTRODUCTION

Patients suffering from the sleep disorder idiopathic rapid-eye-movement sleep behavior disorder (iRBD) are at high risk of developing Parkinson's disease (PD) [1]. In consequence, some studies focus on sleep data in the search for PD biomarkers, where polysomnographic (PSG) data are analyzed either manually or automatic [2] [3]. Supportively, many different attempts to automatic score sleep stages, both in control subjects as well as in sleep disorder patients, have been developed [4] [5]. In [4], a data-driven method was developed, where a topic model with five topics were conducted for each subject based on their sleep electroencephalography (EEG). The method was subject-specific, as it was aimed at providing a complementary approach to sleep analysis by presenting each sleep epoch as a mixture of stages. This study raised the idea of developing a data-driven topic model with the aim of using it to analyze and automatic classify control subjects and patients suffering from either iRBD or PD.

Corresponding author: Julie AEC: jaec@elektro.dtu.dk
Contact information: HBDS: hbs@elektro.dtu.dk, HK: henriette_koch@hotmail.com, RF: RFRA0011@regionh.dk, JK: jkem@elektro.dtu.dk, LA: larv@lundbeck.com, SRC: srac@lundbeck.com and PJ: poje@glo.regionh.dk
[a]DTU Electrical Eng., Ørsteds Plads, bldg. 349, DK-2800 Kgs. Lyngby
[b]H. Lundbeck A/S, Ottiliavej 9, DK-2500 Valby
[c]Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Center for Healthy Aging, Faculty of Health Sciences, University of Copenhagen, Glostrup Hospital, DK-2600 Glostrup

During sleep, eye movements (EMs) are among other structures controlled by neurons located in the brain stem, and in [6], it was found that EMs during sleep hold the possibility of being a PD biomarker. In [6], the obtained performance was based on features reflecting EMs as well as features reflecting electromyography (EMG) measured at the EOG site. Also, the features were computed as the means and standard deviations in energy measures across all sleep epochs during a whole night of sleep. They thereby only reflected the overall differences in EMs between control subjects and iRBD/PD patients. In this study, the focus is on EMs alone, and a general data-driven topic model will be developed illustrating the timely distribution of EMs. A topic model is a statistical model revealing "topics" or "themes", which describe the latent structure behind the generation of a collection of documents. Here, a topic model is applied on data desribing EMs during sleep, and each sleep epoch will be represented as a mixture of three different states for EMs. The three states are thought to be related to slow EMs (SEMs), rapid EMs (REMs) and no EMs (NEMs). By applying the topic model on three test groups of ten control subjects, ten iRBD patients and ten PD patients, it will be analyzed how well the EMs from the patients fall into the normal states for EMs during sleep. By extracting three features from the topic models reflecting "certainty", "fragmentation" and "stability", the test subjects will be classified as "control" or "patient" by use of a Naive Bayes (NB) classifier.

In [7] is a general topic model built on EEG developed based on the same training data as in this study. The number of topics were set to five reflecting the five sleep stages stated in 2004 by the American Academy of Sleep Medicine (AASM) [8]. Features were extracted from the topic models obtained from the same test groups as in this study, and thereby the same subjects as in this study were classified. In this way, this study and [7] reveal how well these patient groups can be classified by application of EOG and EEG, respectively.

II. DATA ACQUISITION

Fourty subjects were enrolled in this study. They were all evaluated at the Danish Center for Sleep Medicine at Glostrup Hospital in Denmark, and the evaluation of the patients included PSG, multiple sleep latency test and a comprehensive medical history and medication. The control subjects included have no history of movement disorder, dream enacting behavior or other former diagnosed sleep

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

disorders. The quality of the PSG data was individually evaluated, and recordings were excluded if the analyzed channels were disconnected or continuously contaminated with artifacts. The demographic data for the groups is seen in Table I.

TABLE I
DEMOGRAPHIC DATA FOR THE FOUR SUBJECT GROUPS

Patient groups	Total No.	Male / Female	Age ($\mu \pm \sigma$) [years]
Controls (for train)	10	5 / 5	57.2 ± 8.1
Controls (for test)	10	5 / 5	59.8 ± 8.4
iRBD (for test)	10	8 / 2	59.0 ± 14.2
PD (for test)	10	6 / 4	63.2 ± 8.4

All subjects underwent at least one full night PSG according to AASM standards by use of different amplifier systems, where the lowest anti-aliasing filter cut-off frequency was 70 Hz. The EOG electrodes were placed one cm out and up (left) or down (right) from the outer canthus with reference to the right and left mastoid, respectively. The sampling frequency of the analyzed sleep data was 256 Hz.

III. METHODOLOGY

The overall methodology of this study is presented schematically in Fig. 1. Ten control subjects selected to best match the patient groups in age were used to develop a general topic model. As input to the topic model, features extracted from bandpass filtered EOG signals were given. By use of the general topic model, 30 topic mixture diagrams were obtained from ten control test subjects, ten iRBD test patients and ten PD test patients. Three features were extracted from these mixture diagrams, and by use of a standard NB classifier, the test subjects were classified as being either "patient" or "control". Below follows a more detailed description of the steps seen in Fig. 1

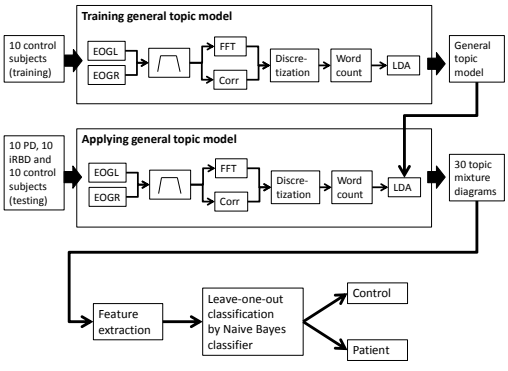


Fig. 1. A schematical overview of the methodology of this study. A general topic model was trained using ten control subjects. The general topic model was applied on ten other control subjects, ten iRBD patients and ten PD patients obtaining 30 topic mixture diagrams. Features were extracted from the topic mixture diagrams, and the subjects were classified as "control" or "patient" using an NB classifier.

A. Generating topic model

Initially, both EOG signals were bandpass filtered by a 4th order Butterworth filter with cut-off frequencies (3 dB) at 0.3 Hz and 10 Hz. These cut-off frequencies were chosen to focus the topic model on EMs by suppressing the influence of the baseline drift, the EMG activity as well as some EEG activity measured at the EOG sites. Both EOG signals were divided into non-overlapping segments of length L, and for each of these segments, three features were computed, yielding a feature vector $f(n)$ expressed as,

$$f(n) = \begin{bmatrix} S_{ll}(n) \\ S_{rr}(n) \\ R_{lr}(n) \end{bmatrix} \quad (1)$$

where n denotes the segment index, S_{ll} and S_{rr} represents the spectral power computed by the fast Fourier Transform (FFT) below 5 Hz in the left and right EOG signal segment, respectively. Any EMs, whether it be SEMs, REMs or a combination of the two, are assumed to be in the range of 0-5 Hz [9]. The R_{lr} represents the normalized cross-correlation coefficient between the left and right EOG signal segment given by,

$$R_{lr}(n) = \frac{\sigma_{lr}(n)}{\sqrt{\sigma_{ll}(n)\sigma_{rr}(n)}} \quad (2)$$

where σ_{ll} and σ_{rr} denotes the variance of the left and right EOG signal segment, respectively, and σ_{lr} denotes the covariance of the left and right EOG signal segment. As the EOG signals appear anticorrelated during EMs, it is assumed that R_{lr} will obtain negative values when REMs occur during REM sleep or wakefulness and when SEMs occur during N1 sleep. Background EOG should appear almost uncorrelated, and the high-amplitude EEG artifacts which can occur during deep sleep should appear correlated. The subject-specific median of the cross-correlation features was subtracted to align the values around zero.

As in [4], the aim is to train a topic model by use of the Latent Dirichlet Allocation (LDA) model. To be able to use the features as input to such a topic model, the features were discretized on a per-subject basis. The spectral power features were given the values 1 to 4 based on boundaries set at each quartile for the full range of feature values for that specific subject. The cross-correlation features were discretized given values 1 to 4 based on boundaries set at [-0.7, 0, 0.7] for all subjects. These boundaries were set based on trial-and-error of best catching the EMs (at values below -0.7), and the EEG artifacts (values above 0.7) as well as the idea of having symmetric boundaries around zero.

The LDA method assumes that a "collection of documents" is derived from an underlying set of "topics", and that the topics are defined as a set of related "words" [10]. As the discretization in this study was done by symbols of 1 to 4, a word length of W is presented by either one of all combinations of W succeeding values of 1 to 4. The LDA assumes that each topic can be defined as a certain distribution over all of the available words. For each document in the collection of documents, a count is formed

of the number of occurrences of each word, and as an end result a topic-by-document matrix X is found, descriping the distribution over topics in each document [10].

As in [4], the document length in this study was set to 30 seconds (comparable with a sleep epoch), yielding that each sleep epoch consisted of a total of $3 \times \frac{30}{L}$ symbol instances. Different word lengths were tried ($W = 2,3,5$), giving that the total number of available words was 3×4^W. The number of topics was set to $T = 3$, in trying to reflect the different states (SEMs, REMs and NEMs) for EMs during sleep.

To train a general topic model, all the available sleep epochs in between lights off and lights on from ten control subjects were used as the collection of documents. By using data from control subjects only, a general "control topic model" was thereby trained. The topic model was applied on the three test groups (see Table I), yielding a topic mixture diagram X holding the distribution of the three "control topics" in each sleep epoch from each of the subjects in the test data.

B. Feature extraction and classification

The aim of this study is to classify the 30 test subjects into either "control" or "patient" based on the topic mixture diagrams obtained when using a general topic model. For each test subject, three features were computed. The features reflect "certainty", "fragmentation" and "stability", and are defined as:

Feature 1 - "Certainty": The amount of epochs with a dominating topic of a probability higher than a given threshold. Normalization was done by dividing the number with the subject-specific total number of epochs. Feature 1 is expressed as,

$$f_1^p = \frac{\sum_{k=1}^{K} logical\left(max(X_k^p) > th\right)}{K} \tag{3}$$

where K is the subject-specific total number of epochs and X_k^p is the EM topic mixture for epoch k in subject p. The threshold value th was defined as the one giving the highest mean Area Under Curve (AUC) when classifying the 30 test subjects using the leave-one-subject-out validation scheme.

Feature 2 - "Fragmentation": The amount of state shifts between topics when the dominating topic defines the state of an epoch. Normalization was done by dividing the number with the subject-specific total number of epochs. Feature 2 is expressed as,

$$f_2^p = \frac{\sum_{k=1}^{K-1} logical\left(max(X_k^p) \neq max(X_{k+1}^p)\right)}{K} \tag{4}$$

Feature 3 - "Stability": The normalized mean number of epochs kept in a certain state when the dominating topic defines the state of an epoch. Feature 3 is expressed as,

$$f_3^p = \frac{\sum_{m=1}^{M} e_m^{new}}{M} \quad \text{with} \quad e^{new} = \frac{e^{old} - \min(e^{old})}{\max(e^{old}) - \min(e^{old})} \tag{5}$$

where m is an index for a period, in where the epochs all have the same dominating topic, M is the subject-specific total number of such periods and e^{old} is a vector holding the M non-normalized numbers of epochs in each period.

As the topic mixture diagrams depend on the initialization of the LDA method, and as it was noticed that the feature values therefore slightly differed in between different runs on the same test subject, the three described features were computed for 20 different runs on the testdata. The mean of the 20 feature values were used as the final feature values. Using the leave-one-subject-out approach, a standard NB classifier was used to classify the subjects into either "control" or "patient". The classification were performed using all combinations of either one, two or all three feature values.

As mentioned earlier, different values were tried for the word length W ($W = 2,3,5$) and for the segment length L ($L = 1,3$). The final topic model developed from the training dataset was chosen based on how well the NB classifier performed (according to accuracy) on the test dataset.

IV. RESULTS AND DISCUSSION

A. Interpreting the topic mixture diagrams

Fig. 2, 3 and 4 present an example of a topic mixture diagram from a control subject, an iRBD patient and a PD patient, respectively. Each vertical coloured bin presents a sleep epoch, and the amount of each color in a bin presents the individual topic probability. Remembering that the three topics are derived based on features reflecting EMs, it is seen, that the general topic model do recognize the characteristic temporal evolution of sleep. More specifically, the "blue" topic could be interpreted as having something to do with the REMs in REM sleep, whereas the "green" topic could be linked to SEMs and the "red" topic could be linked to NEMs. It is seen from the mixture topic diagrams in Fig. 3 and 4, that not as many sleep epochs show a high certainty of either topic as compared to the control mixture diagram in Fig. 2. Interpreting the topics as just described, this observation lead to the conception that the EMs (both the REMs and SEMs) in the patients are less pronounced or less alike the EMs in control subjects. Other observations include the more abrubt transitions in between topics as well as the less structured and more fragmented profiles for the iRBD and PD patients compared to the control subjects. These observations are tried captured in the features "certainty", "fragmentation" and "stability".

B. Classification

A standard NB classifier was used to classify the subjects by the leave-one-subject-out validation approach, and it was found that the model, which obtained the highest mean accuracy, had a segment length of $L = 1$ and a word length of $W = 3$. This model used the features "certainty" and "stability", and in Fig. 5 the decision boundary is illustrated by the colors gray (classified as "patient") and white (classified as "control"). The test subjects are marked by red (PD patient), green (iRBD patient) or blue (control subject) filled circles. It is seen that two control subjects and one iRBD patient are misclassified, yielding a sensitivity of 95 %, a specificity of 80 % and an accuracy of 90 %.

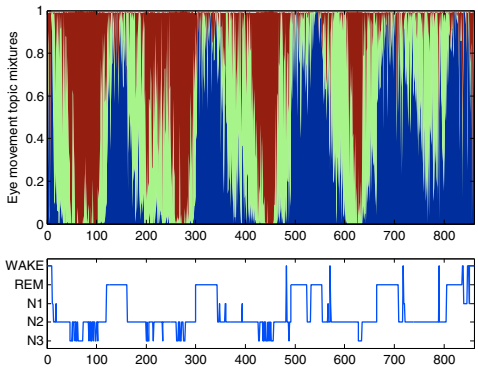


Fig. 2. A topic mixture diagram and the manually scored hypnogram for a control subject.

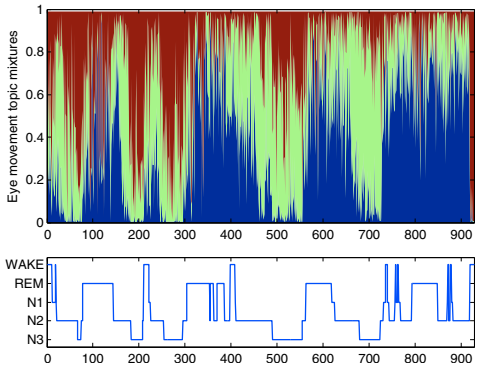


Fig. 3. A topic mixture diagram and the manually scored hypnogram for an iRBD patient.

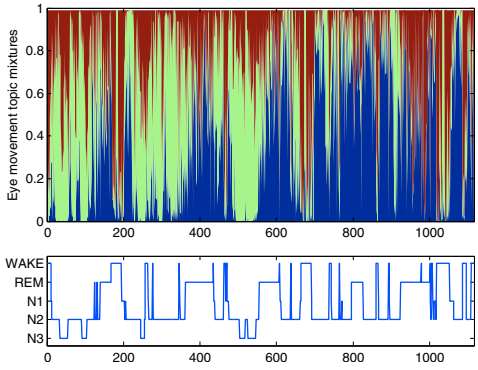


Fig. 4. A topic mixture diagram and the manually scored hypnogram for a PD patient.

V. CONCLUSIONS

Training a general topic model based on sleep EOG from ten control subjects, revealed that the characteristic sleep cycles can be encompassed solely by use of features reflecting EMs. By applying the topic model on testdata from ten other control subjects, ten iRBD patients and ten PD patients, a topic mixture diagram was obtained for each subject. Features reflecting "certainty", "fragmentation" and "stability" of these diagrams were derived. It was found that by use of the features "certainty" and "stability", a simple NB classifier classified the subjects with a sensitivity of 95 %, a specificity of 80 % and an accuracy of 90 %. The separability of the individual features as well as new features derived from the topic mixture diagrams should be further investigated. Although more focused analyzes of the morphology of EMs are needed, this study demonstrates with a data-driven, unsupervised approach that PD and iRBD patients reflect abnorm form and/or timely distribution of EMs during sleep.

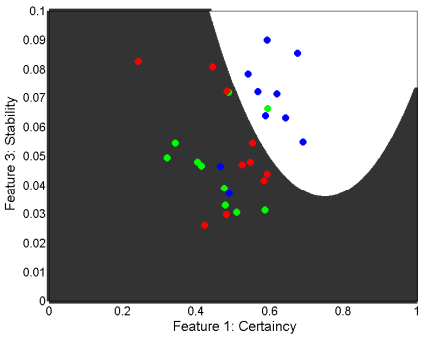


Fig. 5. The best NB classification result was based on two features. The decision boundary is illustrated by the colors white (control area) and gray (patient area), and the 30 test subjects are marked with blue (control subject), green (iRBD patient) or red (PD patient) filled circles.

REFERENCES

[1] C. Schenck, S. Bundlie, and M. Mahowald, "Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behaviour disorder," *Neurology*, vol. 46, no. 2, pp. 388–393, 1996.

[2] R. B. Postuma, J. F. Gagnon, S. Rompré, and J. Y. Montplaisir, "Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease," *Neurology*, vol. 74, pp. 239–244, Jan. 2010.

[3] V. Latreille, J. Carrier, J. Montplaisir, M. Lafortune, and J.-F. Gagnon, "Non-rapid eye movement sleep characteristics in idiopathic REM sleep behavior disorder," *Journal of the Neurological Sciences*, vol. 310, pp. 159–162, June 2011.

[4] A. Van Esbroeck and B. Westover, "Data-driven modeling of sleep states from EEG," in *34th Annual International Conference of the IEEE EMBS*, pp. 5090–5093, 2012.

[5] P. S. Jensen, H. B. D. Sorensen, H. L. Leonthin, and P. Jennum, "Automatic sleep scoring in normals and in individuals with neurodegenerative disorders according to new international sleep scoring criteria," *Journal of Clinical Neurophysiology*, vol. 27, pp. 296–302, Aug. 2010.

[6] J. A. E. Christensen, R. Frandsen, J. Kempfner, L. Arvastson, S. R. Christensen, P. Jennum, and H. B. D. Sorensen, "Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel," in *34th Annual International Conference of the IEEE EMBS*, pp. 2941–2944, 2012.

[7] H. Koch, J. A. E. Christensen, L. Arvastson, S. R. Christensen, P. Jennum, and H. B. D. Sorensen, "Classification of iRBD and Parkinson's patients using a general data-driven sleep staging model built on EEG," in *35th Annual International Conference of the IEEE EMBS [accepted]*, 2013.

[8] C. Iber, *The AASM Manual for the Scoring of Sleep and Associated Events*. American Academy of Sleep Medicine, 2007.

[9] R. Agarwal, T. Takeuchi, S. Laroche, and J. Gotman, "Detection of rapid-eye movements in sleep studies," *IEEE Transactions on Bio-Medical Engineering*, vol. 52, pp. 1390–1396, Aug. 2005.

[10] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

Paper III

TITLE:
Data-driven modeling of sleep EEG and EOG reveals characteristics indicative of pre-Parkinson and Parkinson's disease

AUTHORS:
Julie A. E. Christensen, Marielle Zoetmulder, Henriette Koch, Rune Frandsen, Lars Arvastson, Søren R. Christensen, Poul Jennum and Helge B. D. Sorensen

JOURNAL:
Journal of Neuroscience Methods

YEAR:
2014

VOLUME:
235

PAGES:
262-76

PUBLICATION HISTORY:
Submission date: 25 January 2014, acceptance date: 23 July 2014

Journal of Neuroscience Methods 235 (2014) 262–276

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Clinical Neuroscience

Data-driven modeling of sleep EEG and EOG reveals characteristics indicative of pre-Parkinson's and Parkinson's disease

Julie A.E. Christensen[a,b,d,*], Marielle Zoetmulder[b,e], Henriette Koch[a], Rune Frandsen[b], Lars Arvastson[d], Søren R. Christensen[d], Poul Jennum[b,c], Helge B.D. Sorensen[a]

[a] Department of Electrical Engineering, Technical University of Denmark, Orsteds Plads, Building 349, DK-2800 Kongens Lyngby, Denmark
[b] Danish Center for Sleep Medicine, University of Copenhagen, Department of Clinical Neurophysiology, Glostrup Hospital, Entrance 5, Nordre Ringvej 57, DK-2600 Glostrup, Denmark
[c] Center for Healthy Ageing, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen N, Denmark
[d] H. Lundbeck A/S, Ottiliavej 9, DK-2500 Valby, Denmark
[e] Department of Neurology, Bispebjerg Hospital, Bispebjerg Bakke 23, DK-2400 Copenhagen, Denmark

HIGHLIGHTS

- A data-driven topic modeling approach characterizing sleep EEG and EOG is proposed.
- The approach showed potential for evaluating patients with neurodegeneration.
- The number of topics linked with REM and N3 could be an early PD biomarker.
- The ability to maintain NREM and REM sleep could be an early PD biomarker.
- Patients were classified with 91.4% sensitivity and 68.8% specificity.

ARTICLE INFO

Article history:
Received 25 January 2014
Received in revised form 29 June 2014
Accepted 23 July 2014
Available online 1 August 2014

Keywords:
Parkinson's disease
Topic modeling
Electroencephalography
Electrooculography
Polysomnography
Automatic classification

ABSTRACT

Background: Manual scoring of sleep relies on identifying certain characteristics in polysomnograph (PSG) signals. However, these characteristics are disrupted in patients with neurodegenerative diseases.
New method: This study evaluates sleep using a topic modeling and unsupervised learning approach to identify sleep topics directly from electroencephalography (EEG) and electrooculography (EOG). PSG data from control subjects were used to develop an EOG and an EEG topic model. The models were applied to PSG data from 23 control subjects, 25 patients with periodic leg movements (PLMs), 31 patients with idiopathic REM sleep behavior disorder (iRBD) and 36 patients with Parkinson's disease (PD). The data were divided into training and validation datasets and features reflecting EEG and EOG characteristics based on topics were computed. The most discriminative feature subset for separating iRBD/PD and PLM/controls was estimated using a Lasso-regularized regression model.
Results: The features with highest discriminability were the number and stability of EEG topics linked to REM and N3, respectively. Validation of the model indicated a sensitivity of 91.4% and a specificity of 68.8% when classifying iRBD/PD patients.
Comparison with existing method: The topics showed visual accordance with the manually scored sleep stages, and the features revealed sleep characteristics containing information indicative of neurodegeneration.
Conclusions: This study suggests that the amount of N3 and the ability to maintain NREM and REM sleep have potential as early PD biomarkers. Data-driven analysis of sleep may contribute to the evaluation of neurodegenerative patients.

1. Introduction

Patients suffering from the sleep disorder idiopathic rapid eye-movement sleep behavior disorder (iRBD) have been observed to be at high risk of developing Parkinson's disease (PD) (Iranzo et al.,

* Corresponding author at: Technical University of Denmark, Orsteds Plads, Building 349, DK-2800 Kongens Lyngby, Denmark. Tel.: +45 40255689/45255737.
E-mail address: julie.a.e.christensen@gmail.com (J.A.E. Christensen).

http://dx.doi.org/10.1016/j.jneumeth.2014.07.014
0165-0270/

2006; Schenck et al., 2013, 1996). These longitudinal studies indicate that after a period of approximately 13 years 45% of the patients initially diagnosed with iRBD develop PD or another neurological disorder. Sixteen years after an iRBD diagnosis the incidence rate is 81%. This indicates long latencies from the reported onset of iRBD to the appearance of detectable neurodegeneration. Braak et al. (2003) described the evolution of PD as initially involving the brain stem, then pursuing an ascending course to additional brain areas including the cortex. During this process, there is symptom progression that can potentially be detected in features that are expressed electrophysiologically during sleep. It follows that investigating polysomnograph (PSG) data either manually or automatically may be useful for developing specific and objective markers for PD diagnosis (Christensen et al., 2014; Latreille et al., 2011; Postuma et al., 2010a). According to the American Academy of Sleep Medicine (AASM) Staging Manual from 2007, sleep annotation is done by manually assigning periods of 30 s to either wakefulness (W), rapid eye movement (REM) sleep stage or one of three non-REM sleep stages (N1–N3). Manual scoring relies on identifying certain characteristics in the electrophysiological PSG signals, including the various EEG frequency bands (delta, theta, alpha, beta), EEG microsleep events, such as sleep spindles and K-complexes, and EOG events, such as rapid and slow eye movements (REMs and SEMs, respectively) (Iber et al., 2007). The duration of microsleep events is typically 0.5–3.0 s (Iber et al., 2007).

Reported sleep EEG changes in patients with PD and other neurodegenerative diseases (NDDs) include changed patterns and/or reduced numbers of several sleep-specific phenomena such as sleep spindles (Christensen et al., 2014), changed slow wave characteristics (Latreille et al., 2011) and frequency-slowing (Rodrigues Brazète et al., 2013) relative to age-matched control subjects. As eye movements (EMs) are controlled by neurons located in the brain stem, midbrain areas and frontal areas (Carpenter, 2000), it is believed that patients suffering from an NDD can show affected EMs. Several studies have reported impairment of the oculomotor function in patients with PD (Corin et al., 1972; Mosimann et al., 2005). The oculomotor abnormalities include limitation or absence of gaze in various planes, inadequate convergence and impairment in reflexive saccades during wakefulness. Few studies have investigated PSG EOG in patients with iRBD or PD during sleep, and these reported abnormalities in the outlook and the nightly distribution of rapid and slow EMs (Christensen et al., 2013, 2012).

Stage shifts during sleep and the transition from sleep to wakefulness are controlled by switching mechanisms regulated by several neurons mainly located in the brainstem and midbrain areas (Saper et al., 2001, 2010; Schwartz and Roth, 2008). Proposed models describe the transitions between sleep and wakefulness, and those between REM and NREM as flip-flop switches. These switches are mutually dependent and determine the wake-sleep cyclic rhythm (Lu et al., 2006). The sleep–wake flip-flop switch involves the ascending arousal pathway and the sleep-promoting pathway. The ascending arousal pathway involves a branch of cholinergic neurons, which fire rapidly during wakefulness and REM sleep. Further, the ascending pathway involves a branch of noradrenergic, serotoninergic, dopaminergic and histaminergic neurons, which fire rapidly during wakefulness, less rapidly during NREM sleep and almost cease firing during REM sleep (Saper et al., 2010; Schwartz and Roth, 2008). The neurons of the sleep-promoting pathway inhibit the circuits of the ascending arousal system, and the mutually inhibitory relationship can generate rapid transitions between waking and sleeping states. The REM–NREM flip-flop switch involves two mutually inhibitory populations of GABAergic neurons located in the upper pons, which enables it to generate rapid transitions between REM and NREM sleep states. A malfunction or destruction of any of the loops involved in the two flip-flop switches might be observed in the sleep architecture as either unstructured transitions and/or abnormal amount of time spent in the individual stages. This study investigates whether the neurodegeneration present in iRBD and PD patients affects these mechanisms to a degree that can be revealed by analyzing EEG and EOG. The putative changes in the appearance of the EEG or EOG are linked to the ascending cortical branch of neurons active during REM sleep. It has been shown that the normal descending inhibition of the skeleton muscles in iRBD patients during REM sleep is affected by neurodegeneration (Kempfner et al., 2010; Postuma et al., 2010b). The electromyograph (EMG) activity in these patients is enhanced, which suggests that a neurodegenerative process occurs in the region of the brain stem generating REM atonia (Brown et al., 2012). By not including EMG features, this study focus on the characteristics affected by the ascending cortical parts of the sleep–wake and the REM–NREM sleep switches.

The aim of this study was to evaluate the diagnostic value of features reflecting sleep characteristics such as the stability, fragmentation and distribution of sleep stages in patients with iRBD or PD. To solve the well-known problems with the manual scoring of sleep, we characterized sleep using a data-driven approach based on certain clues from either EEG or EOG. By addressing sleep analysis with a topic modeling approach, we did not attempt to match the manually scored sleep stages. We set out to identify topics in the EEG and EOG and thereby capture latent diversities between subjects with and without neurodegeneration. Data from control subjects were used to develop models defining EEG or EOG topics. To analyze how well the EEG and EOG sleep structures from NDD patients fell into the standard topics, the models were applied to PSG data from additional control subjects, patients with a motor disturbance but not an NDD (periodic leg movements (PLMs)), iRBD patients and PD patients. By extracting features reflecting characteristics of the EEG and EOG topic patterns, this study attempted to reveal sleep characteristics indicative of early and mature neurodegeneration in PD patients.

2. Materials and methods

2.1. Subjects

Subjects were recruited from patients evaluated at the Danish Center for Sleep Medicine (DCSM) in the Department of Clinical Neurophysiology, Glostrup University Hospital. All patient evaluations included a comprehensive medical and medication history. All patients were assessed by PSG and with a multiple sleep latency test (MSLT). Patients treated with medication known to affect sleep stages (antidepressants, antipsychotics, hypnotics) were excluded, although dopaminergic treatments were continued. A total of 36 PD, 31 iRBD and 25 PLM patients were included. The iRBD patients included expressed dream enactment coinciding with REM sleep without atonia (RSWA), manifested as sustained muscle activity in the chin EMG and/or excessive transient muscle activity in either the chin or limb EMG. They were diagnosed with idiopathic RBD since the disease could not be linked to the presence of narcolepsy, neurodegenerative, cerebrovascular, or other neurological disorders or other factors such as specific drugs or psychological state. The PLM patients included did not show any signs of neurodegeneration or RSWA. No PSG findings or NDD-related symptoms were reported for the PLM group, and they were considered as solely PLM patients. Henceforth, the term 'PLM patients' refers to patients suffering solely from PLMs. Thirty-three age-matched control subjects with no history of movement disorder, dream-enacting behavior or other previously diagnosed sleep disorders were included.

The data were split into three groups: one for developing the topic models (10 control subjects), another for training a statistical model for classifying NDD patients (16 subjects from each group),

Table 1
Demographic characteristics of the four groups used to develop, train and validate the sleep models. Means and standard deviations are indicated with μ and σ, respectively.

Patient group	Total counts (male/female)	Age [years, $\mu \pm \sigma$]	BMI [kg/m^2, $\mu \pm \sigma$]	Sleep efficiency [%, $\mu \pm \sigma$]	Time in bed [min, $\mu \pm \sigma$]	LM index	Disease duration [years, $\mu \pm \sigma$]
Data used for building sleep models							
Controls	10 (5/5)	56.0 ± 8.4	23.8 ± 2.7	89.4 ± 5.5	480 ± 52.9	38.6 ± 43.8	NA
Training dataset							
Controls	16 (5/11)	56.7 ± 10.2	23.1 ± 2.6	87.7 ± 8.3	490 ± 72.3	21.6 ± 13.7	NA
PLM	16 (9/7)	56.9 ± 11.8	25.1 ± 3.5	86.5 ± 6.1	434 ± 32.7	60.4 ± 42.3	NA
iRBD	16 (13/3)	62.9 ± 8.6	25.8 ± 3.3	87.4 ± 5.7	485 ± 104.2	44.4 ± 32.0	NA
PD	16 (11/5)	65.4 ± 6.4	26.4 ± 2.6	78.5 ± 13.5	462 ± 87.9	47.1 ± 47.6	3.6 ± 4.5
Validation dataset							
Controls	7 (2/5)	56.9 ± 7.4	23.1 ± 2.1	82.7 ± 15.5	470 ± 103.6	10.8 ± 8.5	NA
PLM	9 (4/5)	57.0 ± 12.1	26.7 ± 4.3	85.0 ± 9.8	428 ± 73.9	49.7 ± 25.1	NA
iRBD	15 (13/2)	63.4 ± 5.8	25.4 ± 3.1	80.8 ± 8.6	486 ± 94.0	32.8 ± 26.4	NA
PD	20 (13/7)	65.1 ± 6.9	25.0 ± 3.4	76.3 ± 13.7	430 ± 66.6	29.4 ± 30.3	8.8 ± 3.5

and a third for the final validation of the classification (7 control subjects, and 9 PLM, 15 iRBD and 20 PD patients). The datasets were designed to be matched by age with no further knowledge of the subjects, except for the building dataset, which was also balanced by gender. The building dataset included 10 control subjects because two pilot studies concerning topic modeling of sleep have shown this number to produce sufficient data (Christensen et al., 2013; Koch et al., 2013). The training and validation dataset sizes were selected on the basis of two criteria: (1) the training dataset should include equal numbers of subjects from the four groups, and (2) the number of subjects in any group in the validation dataset should at least be approximately half that of the corresponding group in the training dataset. Two of the PD patients included in the training dataset were later diagnosed with multiple system atrophy and Lewy body dementia, respectively. The demographic data for the three datasets are summarized in Table 1.

2.2. Polysomnograph recordings

All controls underwent at least one night of PSG recording as outpatients, and all patients underwent at least one night of PSG recording, either as outpatients or in hospital in accordance with the AASM standard (Iber et al., 2007). As a minimum requirement, the PSGs used in this study included 16 channels, comprising EEG (F3, F4, C3, C4, O1, O2) and EOG (EOGL, EOGR) with reference to the mastoids (A1, A2), EMG (TIBL, TIBR, CHIN) and ECG in a referential montage. Two flexible belts measuring abdominal and chest respiratory effort, a finger censor measuring SpO2 and HR and a nasal cannula were also used. PSG was repeated only if the quality of traces was inadequate for clinical use. The quality of the PSG data was individually evaluated, and recordings were excluded if the analyzed channels were disconnected or continuously contaminated with artifacts. Table 2 summarizes the manually scored sleep stage distributions.

The raw sleep data, hypnograms and sleep events were extracted from Somnologica Studio (V5.1, Embla, Broomfield, CO, USA) or Nervus (V5.5, Cephalon DK, Nørresundby, Denmark), using the built-in export data tool. For analysis, data were imported into MATLAB (R2012a, MathWorks, Inc., Natick, MA, USA) and analyzed with a sampling frequency of 256 Hz.

2.3. Data-driven staging of EEG and EOG

The overall method of this study is presented in Fig. 1. Sleep data from 10 control subjects were used to build two general topic models; one based solely on EOG and one based solely on EEG. The general topic models were used to generate two topic mixture diagrams (referred to henceforth in this paper as the EOG or EEG topic diagrams) for the remaining 115 subjects. Based on these diagrams, nine and 13 features were extracted from the EOG and EEG topic diagrams, respectively. All features reflect certain characteristics of the sleep EOG and EEG, described in Section 2.4. The discriminative nature of the features and a subgroup of features was analyzed for the two-class problem "NDD" versus "non-NDD" by training a regularized logistic regression model. A detailed description of the steps in Fig. 1 is provided below.

2.3.1. Generating topic models

Topic models were first used to investigate topics in a collection of text documents (Blei et al., 2003). In this context, a topic model is a statistical model that reveals underlying themes or topics describing the latent structures behind the generation of a collection of text documents. Given that each document concerns a given topic, it may be assumed that particular words in the document

Table 2
The distribution of sleep stages used in the development, training and validation of the EEG and EOG sleep models.

Patient group	Wake [number of epochs (%)]	REM [number of epochs (%)]	N1 [number of epochs (%)]	N2 [number of epochs (%)]	N3 [number of epochs (%)]	Sum [number of epochs (%)]
Data used for building sleep models **(epochs were selected to have an approximately equal number in each stage)**						
Controls	642 (20)	700 (22)	617 (19)	700 (22)	597 (18)	3256 (100)
Training dataset						
Controls	1887 (12)	3114 (20)	1292 (8)	7091 (45)	3201 (15)	15,685 (100)
PLM	1836 (13)	2767 (20)	1090 (8)	6010 (43)	2170 (16)	13,873 (100)
iRBD	2012 (13)	2923 (19)	1637 (11)	6789 (44)	2149 (14)	15,510 (100)
PD	3113 (21)	1652 (11)	1867 (13)	6251 (42)	1889 (13)	14,772 (100)
Validation dataset						
Controls	1137 (17)	1136 (17)	683 (10)	2890 (44)	733 (11)	6579 (100)
PLM	1141 (15)	1204 (16)	732 (10)	3674 (48)	951 (12)	7702 (100)
iRBD	2967 (20)	2098 (14)	1106 (8)	5880 (40)	2535 (17)	14,586 (100)
PD	4057 (24)	2464 (14)	1521 (9)	6523 (38)	2625 (15)	17,190 (100)

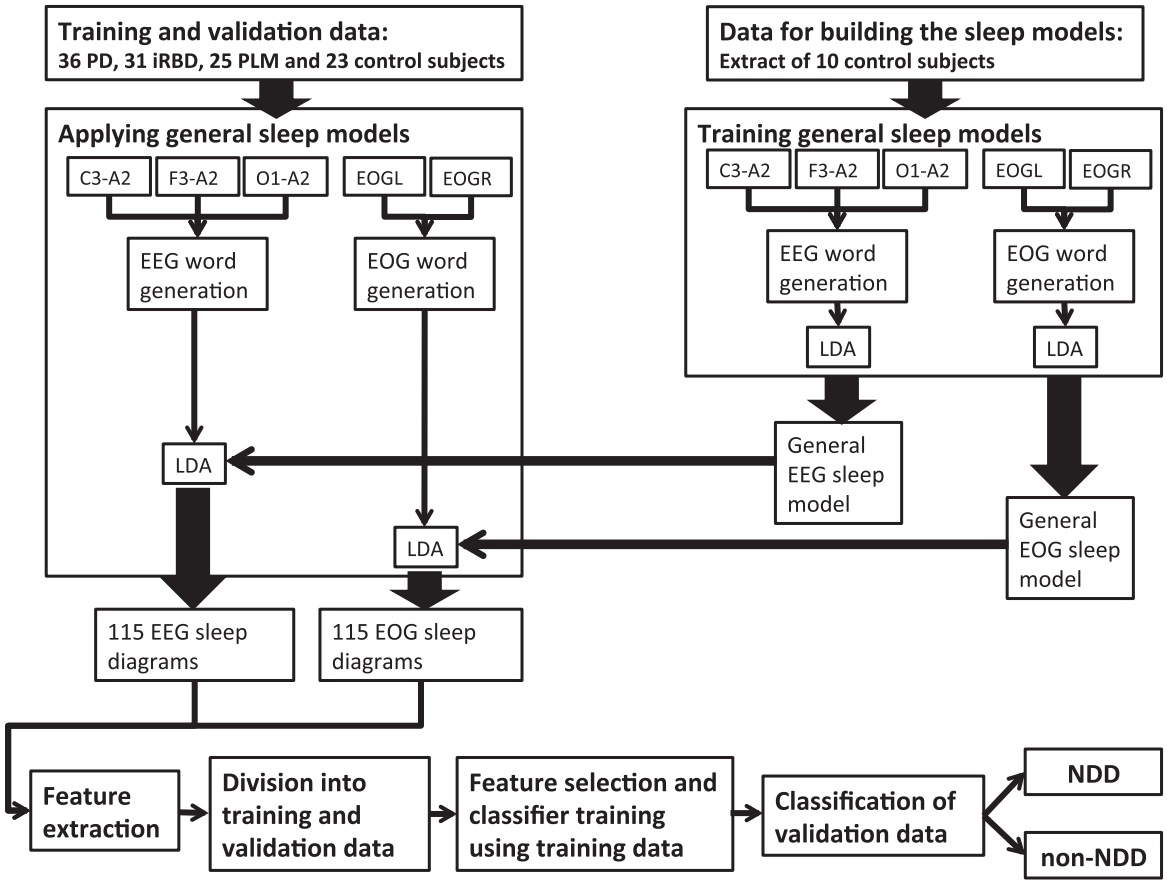


Fig. 1. A schematic overview of the method adopted in this study. An extract of the sleep data from ten control subjects was used to train an EOG topic model using the EOGL-A2 and EOGR-A1 derivations and an EEG topic model using the C3-A2, F3-A2 and O1-A2 derivations. The trained topic models were applied to sleep data from 23 control subjects, and 25 PLM, 31 iRBD and 36 PD patients, yielding two sleep diagrams per subject. Nine features were extracted from the EOG topic diagrams and 13 features were extracted from the EEG topic diagrams. Using a training dataset of 16 subjects from each group, feature subsets were selected and three regularized logistic regression models were trained to classify each subject as either "NDD patient" or "non-NDD subject". The performance of the regression models was assessed in the validation dataset.

are strongly linked to that specific topic. As an example, words such as "computer", "screen" and "keyboard" will be frequent in a document about "IT equipment", whereas words such as "ring", "gold" and "necklace" will be frequent in a document concerning "jewellery". The statistics of the appearances of words in a collection of documents are used to identify what the topics include as well as the topic mixture in each document (Blei et al., 2003).

In the context of sleep data, each sleep epoch is a document. The words used for revealing the topic mixtures can be explained as variables that represent characteristics (EEG and EOG characteristics) of the latent or immanent structures in the different sleep stages. As an example, a "word" linked to a delta wave will appear more frequently in an N3 epoch, whereas "words" linked to sleep spindles or theta waves will appear more frequently in an N2 epoch. In this way, a topic model can find topics and reveal the mixture of topics for each sleep epoch. Intuitively, the topics are linked to the standard sleep stages but topics and standard sleep stages are not identical.

The EEG words were extracted from the C3-A2, F3-A2 and O1-A2 EEG derivations, which initially were forward and reverse bandpass-filtered with a fourth-order Butterworth filter with cut-off frequencies (3 dB) at 0.3 Hz and 35 Hz, in concordance with the AASM standard. The signals were divided into non-overlapping 1-s segments, m, and four word features were computed for each of these segments. This yielded a vector $T_{\mathrm{EEG}}(m)$ expressed as,

$$T_{\mathrm{EEG}}(m) = \begin{bmatrix} X_{\mathrm{C3-A2}}(m) \\ X_{\mathrm{F3-A2}}(m) \\ X_{\mathrm{O1-A2}}(m) \end{bmatrix} \quad \text{where} \quad X(m) = \begin{bmatrix} x_{\delta}(m) \\ x_{\theta}(m) \\ x_{\alpha}(m) \\ x_{\beta}(m) \end{bmatrix} \tag{1}$$

represents the spectral power in the clinical EEG frequency bands delta ($f < 4$ Hz), theta ($4\,\mathrm{Hz} \le f < 8\,\mathrm{Hz}$), alpha ($8\,\mathrm{Hz} \le f < 13\,\mathrm{Hz}$) and beta ($13\,\mathrm{Hz} \le f < 30\,\mathrm{Hz}$). The spectral power was computed by the fast Fourier transform (FFT) using zeropadding and a rectangular window function.

The feature vector $T_{\mathrm{EEG}}(m)$ was converted into words by discretizing the spectral power on a per-subject basis, given the values 1–5 based on boundaries set at each quintile for the full range of features for that specific subject. For each EEG channel, the quintile boundaries were set individually for each spectral power band based on the full range of power values for one subject across the whole night. Fig. 2 illustrates how the EEG words for a frontal EEG channel were generated. The discretized values 1–5 are illustrated as the height of the colored bins and the clinical band is indicated by the color. Given a word length of three, an EEG word is presented

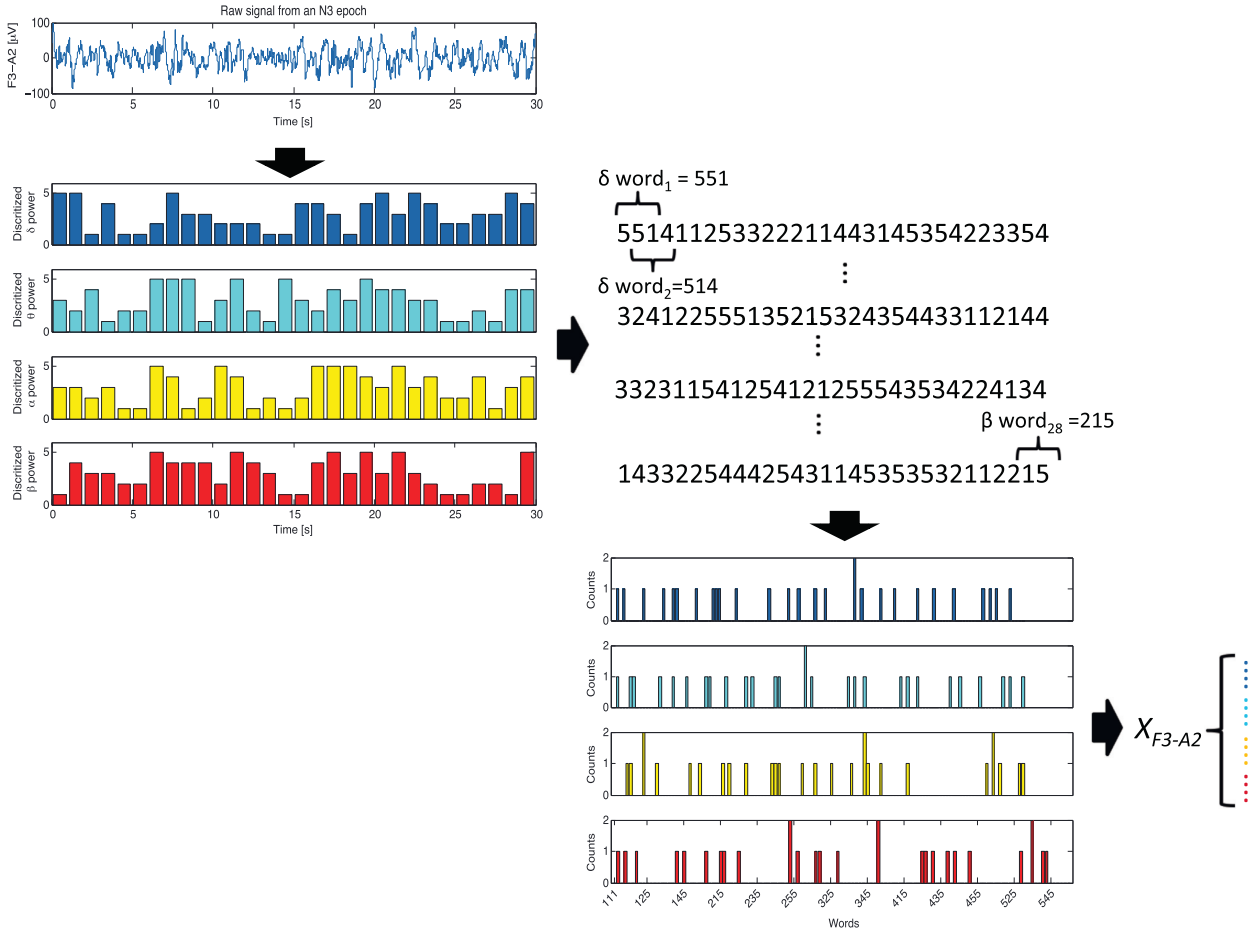


Fig. 2. Illustration of how the EEG words were generated for a frontal EEG signal in a single epoch. For each second, the delta, theta, alpha and beta powers were computed and discretized to obtain values between 1 and 5. The EEG words were defined as any combination of three consecutive values of 1–5 and the final output was a vector of the word counts across the four power bands. Similar vectors were obtained for the central and occipital EEG signal, and combined they make up a specific fingerprint of the sleep epoch analyzed.

as either one of all combinations of three succeeding values of 1–5. The length of three was chosen to reflect the typical length of microsleep events, and the range of five was chosen, as successful results have previously been reported using this length (Koch et al., 2013; Van Esbroeck and Westover, 2012). The available words in the three EEG channels were counted for each sleep epoch, and the final fingerprint of the given sleep epoch consisted of a distribution over all possible words across the three different EEG channels. Specifically, the final EEG fingerprint for a given epoch was a column vector of word distributions from the central, frontal and occipital EEG channels (Fig. 2 shows solely the frontal channel).

The EOG words were extracted from the EOGR-A1 and EOGL-A2 derivations, which were forward and reverse bandpass-filtered with a fourth-order Butterworth filter with cutoff frequencies (3 dB) at 0.3 Hz and 10 Hz. These cutoff frequencies were chosen to focus the analysis on eye movements by suppressing the influence of baseline drift, EMG activity and some EEG activity measured at the EOG sites. For each segment, three word features were computed, yielding a word feature vector $T_{\mathrm{EOG}}(m)$ expressed by

$$T_{\mathrm{EOG}}(m) = \begin{bmatrix} X_{ll}(m) \\ X_{rr}(m) \\ X_{lr}(m) \end{bmatrix}, \quad \text{where} \quad X_{lr}(m) = \sqrt{\frac{\sigma_{lr}(m)}{\sigma_{ll}^2(m)\sigma_{rr}^2(m)}} \tag{2}$$

represents the normalized cross-correlation coefficient between the left and right EOG signal segments m. σ_{ll}^2 and σ_{rr}^2 denote the variance of the left and right EOG signal segments, respectively, and σ_{lr} denotes the covariance of the left and right EOG signal segments. X_{ll} and X_{rr} represent the spectral power below 5 Hz in the left and right EOG signal segments, respectively, computed by FFT using zeropadding and a rectangular window. All eye movements, whether they were SEMs, REMs or a combination of the two, are assumed to be in the range of 0–5 Hz (Agarwal et al., 2005; Christensen et al., 2013, 2012).

The word feature vector $T_{\mathrm{EOG}}(m)$ was converted into words in a similar manner as for $T_{\mathrm{EEG}}(m)$. The EOG spectral power was assigned values between 1 and 3 based on boundaries set at each tertile. The tertile boundaries for each EOG power feature were established on a subject-specific basis over the full range of power values throughout the entire night. The subject-specific median of the cross-correlation features was subtracted, and the normalized cross-correlation features assigned the values 1–3 on the basis of boundaries set at [-0.7; 0.7] for all subjects. These boundaries were set on a trial-and-error basis for the best separation of the data into either EMs (values below −0.7), background EOG (values between −0.7 and 0.7) or EEG artifacts measured at the EOG site (values above 0.7). The length of the EOG words was set at two, given that an EOG word in this study is presented as either one of all

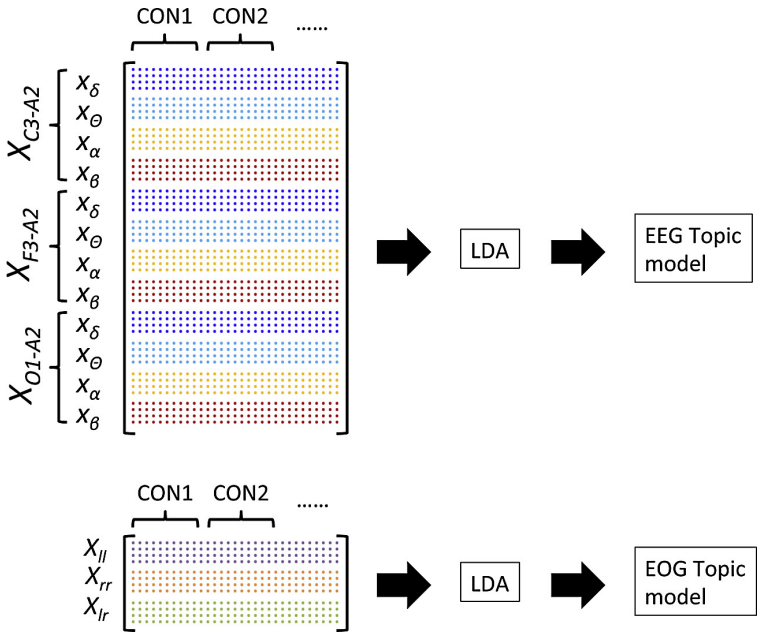


Fig. 3. Illustration of how the vectors describing the word distribution for each epoch conform the input matrices to the LDA. Each column of the matrices presents a sleep epoch, and CON1 and CON2 illustrate different subjects. In the case of EEG, three vectors were combined, containing the word distribution across the central (X_{C3-A2}), frontal (X_{F3-A2}) and occipital (X_{O1-A2}) EEG signals, respectively. Each of these vectors contains the discretized power values for the clinical bands delta (x_δ), theta (x_θ), alpha (x_α) and beta (x_β), as illustrated in Fig. 2. In the case of EOG, two vectors of discretized power values of the left (X_{ll}) and right (X_{rr}) EOG channels, respectively, and a vector of the discretized cross-correlation coefficient measures (X_{lr}) were combined.

combinations of two succeeding values of 1–3. The choice of the boundary values, the length of the EOG words and the range length were all set on the basis of the experience gained and the conclusions of the pilot study described by Christensen et al. (2013). The final EOG fingerprint of an epoch was a column vector of the word distribution of the discretized power measures of the left and right EOG channels and the discretized cross-correlation coefficients.

This study uses Latent Dirichlet Allocation (LDA) (Blei et al., 2003), and the MATLAB-based LDA toolbox implemented by Verbeek (2006). LDA is a common approach used in topic modeling and assumes that a given document contains a combination of multiple topics. This allows identification of concurrent topics and topics, which are related to a small subset of words. Each topic is assumed to be defined by a certain distribution of all available words, and the number of occurrences of each word in each document is counted. In this way, a multinormal distribution over topics is derived for every document. Relating this to sleep analysis, a specific distribution over sleep topics k is provided for every sleep epoch n, and LDA outputs the posterior probability $p_k(n)$ for sleep topic k in epoch n.

Ten control subjects were used to build the general EEG and EOG topic models. By using data only from control subjects, a general "control EOG topic model" and a "control EEG topic model" were trained using the EOG and EEG words separately. As we aimed to use equal representation of the AASM sleep stages in the building process, 70 sleep epochs from each AASM sleep stage were taken out randomly in between lights-off and lights-on from each of the ten control subjects. In the cases where 70 sleep epochs of a given AASM stage were not present (only the case for N1, N3 or W), all the available epochs of that stage were included. An equal number of epochs from each AASM stage were included to give each stage the same chance of being described by a topic. The final number and distribution of sleep epochs used in the building process are shown in Table 2. A matrix containing the word counts for each of the epochs included in the building dataset was fed into LDA. Fig. 3 illustrates how the matrix was conformed for the EEG and the EOG models, respectively. To reflect the three major EMs during sleep states (SEMs, REMs and no EMs), the number of topics in the EOG model was set to three. The number of topics in the EEG model was set to five to reflect the five sleep stages (N1, N2, N3, REM and W).

The general sleep models were applied in all epochs between lights-off and lights-on from 36 PD, 31 iRBD and 25 PLM patients and 23 control subjects. Only the raw frontal, central and occipital EEG signals and the two raw EOG signals were used as input and no manual staging or any other subjective information was needed. An EEG topic diagram and an EOG topic diagram were obtained for each subject, describing the distribution over EEG or EOG topics in each sleep epoch, respectively. Fig. 4 shows examples of an EOG and an EEG topic mixture diagram from a control subject. Each vertical bin presents a sleep epoch, and the amount of each of the colors in each bin represents the probabilities of each topic. If a bin is dominated by red, for example, this reflects a high probability of the red topic. For the purposes of comparison, the manually scored hypnogram is provided below each topic diagram.

2.4. Feature extraction

The aim of this study was to identify sleep EEG and sleep EOG features or characteristics indicative of neurodegeneration. The approach sought to classify the subjects as "NDD patient", containing iRBD and PD patients, or "non-NDD subject", containing PLM patients and control subjects, based on features extracted from the topic diagrams. A description of the features computed is provided below and the reflections and topic diagram observations that gave rise to the feature ideas are stated. Table 3 provides an overview by stating the feature name, a short explanation and the total number of features computed for each feature group.

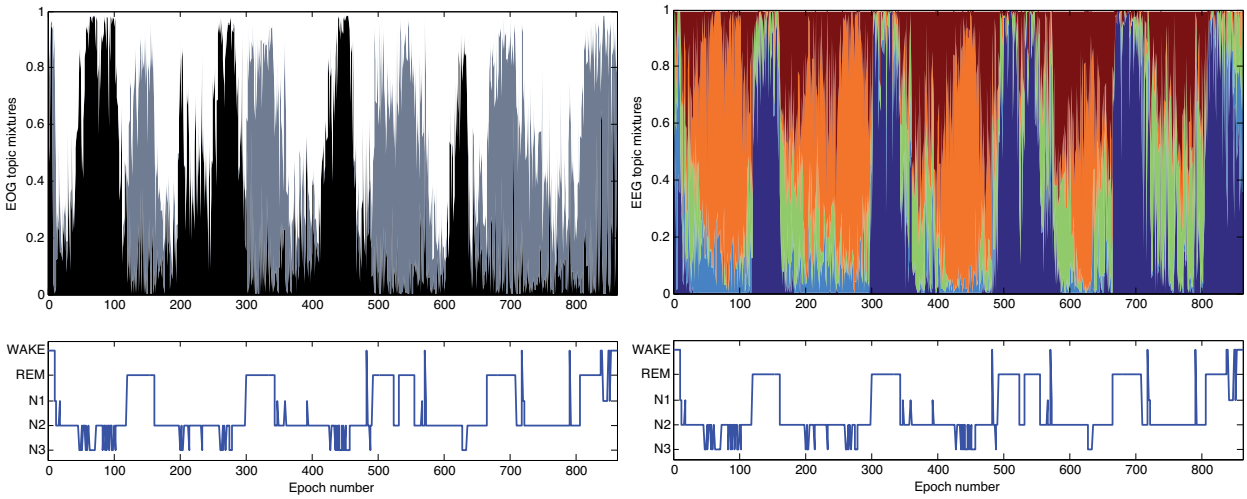


Fig. 4. Examples of an EEG and an EOG topic diagram from a control subject. The figures are stacked percentage column charts, where a sleep epoch is presented as a vertical line made up of a mixture of colors. Each color presents an EOG (black, gray or white) or EEG (dark-blue, light-blue, green, orange or red) topic, where the amount of color in each vertical bin represents the probability of that topic. For comparison, the manually scored hypnogram is presented below the topic diagrams. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4.1. Certainty feature

It was hypothesized that control subjects have a more "clear" sleep outlook, since they show distinct differences between the sleep stages. Comparison of the control/PLM and iRBD/PD topic diagrams indicates that the control subjects and PLM patients have more epochs with high topic probability, i.e., more vertical bins with a high probability of any color. This characteristic was captured in a feature, $f_{\text{certainty}}$, which contains the number of epochs with any given topic probability higher than a given threshold. The feature was normalized with respect to the total number of epochs. $p_k(n)$ defines the probability of topic k in epoch n, thereby the certainty feature can be expressed as:

$$f_{\text{certainty}} = \frac{1}{N}\sum_{n=1}^{N} 1_{\{p_k(n)>t\}} \tag{3}$$

where N is the subject-specific total number of epochs, $t = 0.9$ is a threshold value defining when an epoch is counted as certain and 1 is an indicator function. For each subject, the feature $f_{\text{certainty}}$ was computed for the EOG and EEG topic diagram.

2.4.2. Fragmentation feature

It was hypothesized that iRBD and PD patients show more abrupted sleep with more transitions between the different sleep/EM states. Comparison of the control/PLM and iRBD/PD topic diagrams indicates less structure and more fragmentation in the diagrams of the iRBD and PD patients, for the EOG and specifically for the EEG topic diagrams. This characteristic was captured in a feature, $f_{\text{fragmentation}}$, which encapsulates the number of transitions from one topic to another (from one dominant color to another). The feature was normalized with respect to the total number of epochs N and is expressed as:

$$f_{\text{fragmentation}} = \frac{1}{N}\sum_{n=1}^{N-1} 1_{\{arg\max_k p_k(n) \neq arg\max_k p_k(n+1)\}} \tag{4}$$

The $f_{\text{fragmentation}}$ feature was computed for the EOG and the EEG topic diagrams for each subject.

2.4.3. Global stability feature

It was hypothesized that iRBD and PD patients have a decreased ability to sustain a given sleep/EM state and therefore will stay less time in any sleep stage before switching to another stage. This characteristic is reflected in the topic diagrams as the patients show less structure and few subsequent epochs with the same dominant color. This was captured in a feature, $f_{\text{stability}}$, which encompasses the mean number of epochs in a period s, in which the dominant topic is maintained. A vector l_s expresses the number of epochs in such periods, and is illustrated in Fig. 5. Before computing the mean, l_s was normalized by the subject-specific minimum and maximum values. $f_{\text{stability}}$ is expressed as:

$$f_{\text{stability}} = \frac{1}{S}\sum_{S=1}^{S} L_S \quad \text{with} \quad L_S = \frac{l_S - \min(l_S)}{\max(l_S) - \min(l_S)}. \tag{5}$$

Table 3
Overview of the feature groups computed in this study. The certainty, fragmentation and global stability features were computed once for each topic model. The topic amount feature group and the topic stability feature group each comprise five features computed from the EEG topic diagrams (based on the dark-blue, light-blue, green, orange and red topics) and three features computed from the EOG topic diagrams (based on the black, gray and white topics).

Feature name	Feature explanation	Total number per subject	
		EEG	EOG
Certainty	Normalized number of epochs with any topic probability higher than 90%	1	1
Fragmentation	Normalized number of transitions from one topic to another topic	1	1
Global stability	Normalized number of epochs in a period of epochs where the topic is the same	1	1
Topic amount	Normalized number of epochs with a given topic, and with a topic probability higher than 70%	5	3
Topic stability	Normalized number of epochs in a period of epochs with the same given topic	5	3

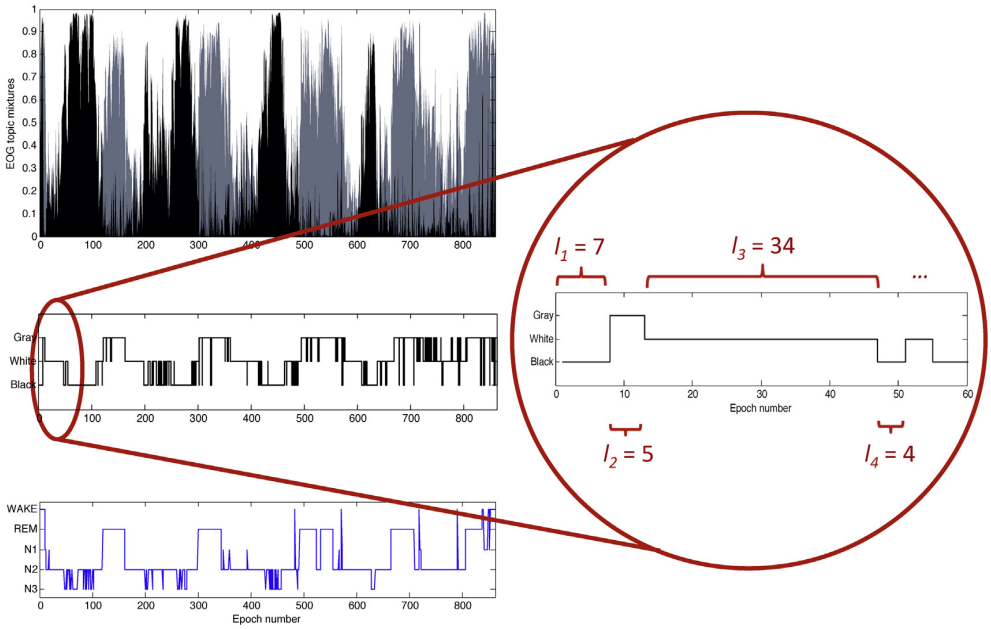


Fig. 5. Illustration of how the vector l used in the stability features described in Section 2.4.3 is computed. A given topic diagram with mixtures of topics in each bin is converted to a diagram, in which the dominant topic defines the epoch. The number of epochs in each period for which the same topic is dominant is computed and forms the elements in vector l.

The feature was computed for the EOG and the EEG topic diagrams for each subject.

2.4.4. Topic amount feature

It was hypothesized that iRBD and PD patients express changed amounts of the different sleep stages during the night. This was captured in the topic amount feature $f^k_{T_{\text{amount}}}$, which was computed for every topic, k, and can be expressed as:

$$f^k_{T_{\text{amount}}} = \frac{1}{N}\sum_{n=1}^{N} 1_{\{p_k(n) > t_k\}}, \tag{6}$$

where k is the topic number and $t_k = 0.7$ is a certainty threshold defining when an epoch is counted as a topic k epoch. The feature $f^k_{T_{\text{amount}}}$ was computed for each of the three EOG topics and five EEG topics for each subject.

2.4.5. Topic stability feature

It was hypothesized that the ability to stay in a given sleep stage is affected by the neurodegeneration in the basal ganglia, which the sleep diagrams support as the patients rarely remain in a given topic for a long time. The overall measure of this characteristic was captured in the feature $f_{\text{stability}}$ described in Section 2.4.3. The ability to stay in specific EEG or EOG topics was not represented by $f_{\text{stability}}$ but was encapsulated by the topic stability feature $f^k_{T_{\text{stability}}}$, which was computed for every topic k. This is expressed as:

$$f^k_{T_{\text{stability}}} = \frac{1}{N}\sum_{n=1}^{N-1} 1_{\{\arg\max_j p_j(n) \neq k \text{ and } \arg\max_j p_j(n+1)=k\}}, \tag{7}$$

and was computed for each of the three EOG topics and five EEG topics for each subject. It reflects the ability of the subject to maintain any of the five EEG topics or the three EOG topics.

2.5. Feature ranking and classification

In total, 22 features were computed for each subject. Nine and 13 reflect characteristics of the EOG and EEG topic diagrams, respectively. The aim of this study was to identify the most indicative features for classifying subjects as "NDD patients" or "non-NDD subjects". Some of the 22 features may be irrelevant or redundant for this classification, and therefore the optimal subset of features was determined using a logistic regression with Lasso regularization. This classification method was chosen because the coefficients are easy to interpret and the regularization forces the coefficients of irrelevant or redundant features to zero, yielding simple models (Hastie et al., 2008; Tibshirani, 1996). Logistic regression is a binary and linear classification method that uses the logistic sigmoid function to model the posterior probability of the *positive* class for the input variable $\mathbf{y}$ (Hastie et al., 2008):

$$p(\text{ positive}\,|\, y) = \frac{1}{1 + e^{-(\beta_0 + \beta^T y)}} \tag{8}$$

where β_0 is the model offset and $\boldsymbol{\beta}$ is a vector of the variable coefficients. The final classification is made by defining the decision boundary $\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x} = 0$ for the logit transformation

$$\log \frac{p(\text{positive}|y)}{1 - p(\text{positive}|\, y)} = \beta_0 + \boldsymbol{\beta}^T y. \tag{9}$$

Maximum likelihood optimization was used for training the model, and the complexity of the model was controlled by introducing the regularization term $\lambda > 0$, yielding an optimization problem expressed as:

$$\underset{(\beta_0+\boldsymbol{\beta})}{\arg\max} \left\{ \sum_{i=1}^{I} \left[\boldsymbol{z}_i(\beta_o + \beta^T \boldsymbol{y}_i) - \log(1 + e^{\beta_0 + \boldsymbol{\beta}^T \boldsymbol{y}_i}) \right] - \lambda \sum_{d=1}^{D} |\boldsymbol{\beta}_d| \right\}, \tag{10}$$

where I is the number of subjects, z_i the classification output stating $z_i = 1$ for positive cases ("NDD") and $z_i = 0$ for negative cases ("non-NDD") and D is the dimension of the input variable. As the

Table 4
Overview of visual interpretation of the topics obtained from the EOG and EEG sleep models.

Topic	EEG "dark-blue"	EEG "light-blue"	EEG "green"	EEG "orange"	EEG "red"	EOG "black"	EOG "gray"	EOG "white"
AASM stage	REM	W	N1/N2	N3	N1/N2	N3	REM/W	N1/N2

regularization term λ increases, the number of nonzero components of $\boldsymbol{\beta}$ decreases, and the model becomes more sparse. Features with nonzero components are included in the final feature subset, and the component values indicate each feature's impact on the final posterior probability output.

The data were split into training and validation datasets, the former consisting of 16 control subjects, and 16 PLM, iRBD and PD patients each. Only this dataset was used for optimizing the classification model, and all input features were standardized. Eight-fold cross-validation was used to optimize the logistic regression models. Three such models were trained: (1) one in which all features were available initially, (2) one in which only the EOG features were available initially and (3) one in which only the EEG features were available initially. The optimal feature subset was found for each regression model, thereby enabling the performance of the EOG and EEG features as well as the combination of the two modalities to be investigated. Performance measures were obtained for the training and validation datasets.

3. Results

3.1. Interpretation of topic models

Fig. 4 shows an example of an EEG and an EOG topic diagram from a control subject. Each sleep epoch is presented as a vertical bin with colors. Each color represents an automatically found topic, and the amount of color in each vertical bin represents the topic probability. The manually scored hypnogram is given below each topic diagram. To discriminate between the EEG and EOG diagrams, the EEG topics are illustrated in bright colors and the EOG topics are shown in black, white or gray.

Clear concordances between the manually scored hypnogram and the EEG topics are seen. There was a link between N3 and the orange topic, REM sleep and the dark-blue topic and W was linked to the light-blue topic. The green and red EEG topics were present in many sleep epochs and therefore more difficult to link to a specific sleep stage. However, they were both pronounced during N1 and N2 sleep. An overview of the visual interpretation of the different EEG and EOG topics is provided in Table 4.

The EOG model was built on words derived from the EOG derivations but it is assumed to recognize the characteristic temporal evolution of sleep. Visual inspection showed that the gray and black EOG topics were most pronounced during REM and N3 sleep, respectively. Specifically, the black EOG topic may be related to low-frequency high-amplitude EEG artifacts measured at the EOG during N3. The gray topic is interpreted as being related to the REMs during REM sleep, whereas the white topic may include both the low-amplitude SEMs and periods with no EMs.

Fig. 6 shows examples of EEG and EOG topic diagrams from PLM, iRBD and PD patients. The diagrams from the iRBD and PD patients had a less structured appearance, more fragmented profiles and more abrupt transitions between topics compared with those from the PLM patients and control subjects in Fig. 4. Further, fewer sleep epochs in the iRBD and PD diagrams had a high certainty of either topic compared with the sleep diagrams from a control subject. These observations and characteristics are captured in the features explained in Section 2.4.

3.2. Feature subsets and classification performances

Fig. 7 presents the feature values extracted from the two topic diagrams from each of the subjects and indicates that some of the features had discriminative properties for iRBD and PD patients versus PLM patients and control subjects. Feature subsets were selected when including all the 22 features in order to investigate the discriminating ability of the features and how they perform together. Additionally, feature subsets were selected when including only the EOG features or only the EEG features. The regression model initially including all 13 EEG features ended up selecting four EEG features, as shown by the logistic regression model obtained:

$$p(\mathrm{NDD}|\boldsymbol{f}_{\mathrm{EEG}}) = \frac{1}{1+\exp(-0.002+0.363f^{\text{Dark Blue}}_{T_{\text{Amount}}}+0.311f^{\text{Orange}}_{T_{\text{Amount}}}-0.310f^{\text{Green}}_{T_{\text{Stability}}}+0.343f^{\text{Orange}}_{T_{\text{Stability}}})}. \quad (11)$$

This expression indicates that small values of $f^{\text{Dark Blue}}_{T_{\text{Amount}}}$, $f^{\text{Orange}}_{T_{\text{Amount}}}$ and $f^{\text{Orange}}_{T_{\text{Stability}}}$ and a large value of $f^{\text{Green}}_{T_{\text{Stability}}}$ increased the probability of belonging to the NDD class. The coefficients show that $f^{\text{Dark Blue}}_{T_{\text{Amount}}}$, followed by $f^{\text{Orange}}_{T_{\text{Stability}}}$, $f^{\text{Orange}}_{T_{\text{Amount}}}$ and $f^{\text{Green}}_{T_{\text{Stability}}}$ were the most indicative. The dark-blue, orange and green topics are interpreted as REM, N3 and N1/N2 sleep, respectively. It was found that the amount of REM sleep, and the amount and stability of N3 were considerably higher for the control/PLM group than for the iRBD/PD group, which is reflected by high positive β values in the regression model. The stability measure of N1/N2 was less for the control/PLM group than for the iRBD/PD group, which is reflected by a negative β value in the regression model.

The regression model initially including all nine EOG features ended up using two EOG features, as shown by the logistic regression model obtained:

$$p(\mathrm{NDD}|\boldsymbol{f}_{\mathrm{EOG}}) = \frac{1}{1+\exp(-0.003+0.011f_{\text{Certainty}}+0.320f^{\text{Black}}_{T_{\text{Amount}}})}. \quad (12)$$

This expression indicates that small values of $f^{\text{Black}}_{T_{\text{Amount}}}$ and $f_{\text{Certainty}}$ increased the probability of belonging to the NDD class. $f^{\text{Black}}_{T_{\text{Amount}}}$ was the most indicative and may be linked to the EEG activity at EOG sites during N3 sleep. The coefficients show that the amount of EEG activity at EOG sites during N3 sleep was considerably higher and the overall EOG certainty measure was marginally higher for the control/PLM group than for the iRBD/PD group.

The regression model initially including all 22 features selected the same four EEG features as the EEG regression model and gave an identical model expression.

Table 5 presents the performance measures for the final regression models and the optimal feature subset for each. Performance measures were computed using the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN), stated as:

$$\text{sensitivity} = \frac{\#\mathrm{TP}}{\#\mathrm{TP}+\#\mathrm{FN}} \quad \text{and} \quad \text{specificity} = \frac{\#\mathrm{TN}}{\#\mathrm{TN}+\#\mathrm{FP}}. \quad (13)$$

A subject is detected as NDD when $p(\mathrm{NDD}|\boldsymbol{f}) \geq 0.5$ and TP and FN represent the number of iRBD and PD patients detected as NDD and

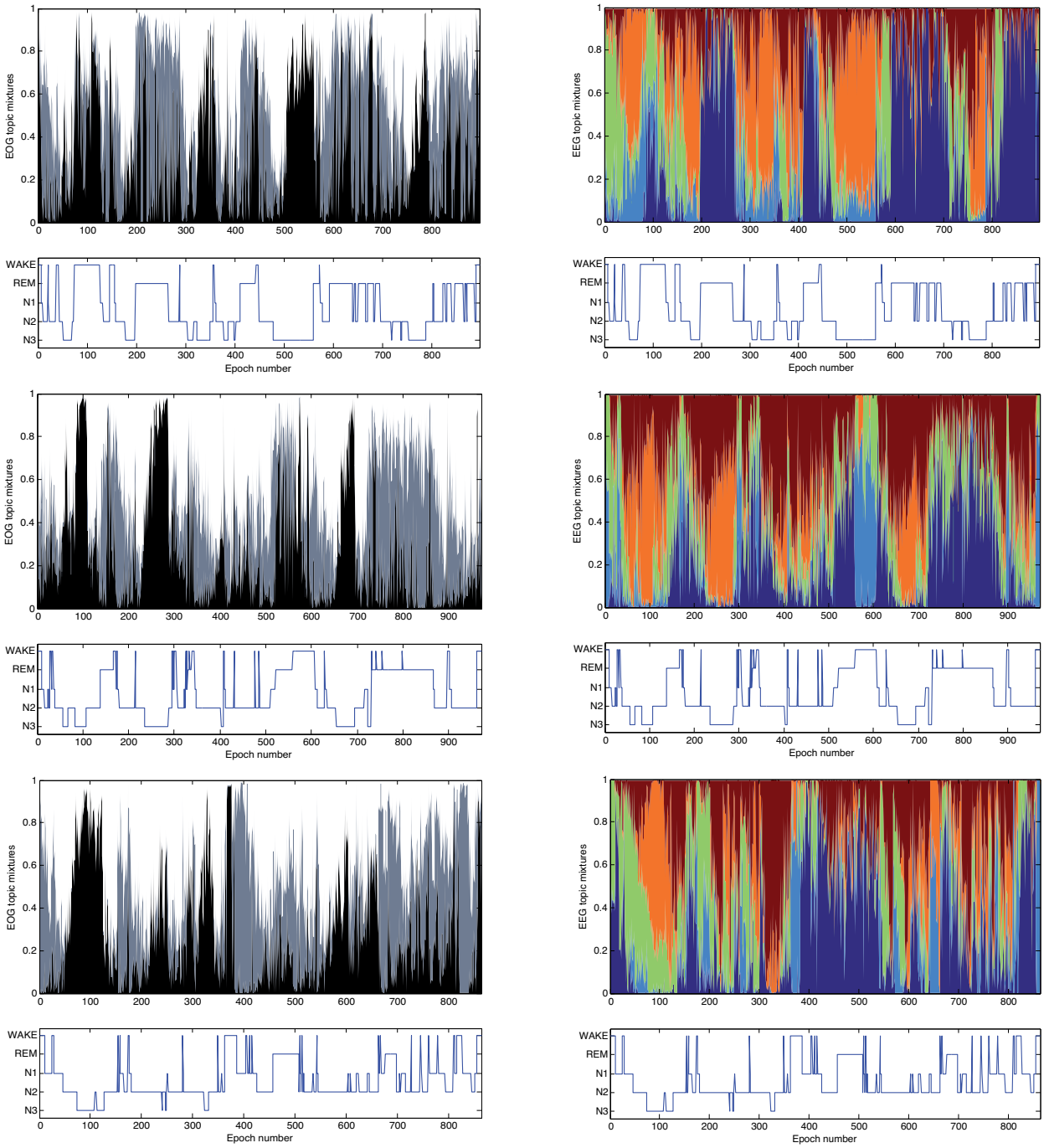


Fig. 6. Examples of EEG and EOG topic diagrams from a PLM (top), an iRBD (middle) and a PD (bottom) patient. The figures are stacked percentage column charts, where a sleep epoch is presented as a vertical line possessing a mixture of colors. Each color presents an EOG (black, gray or white) or EEG (dark blue, light blue, green, orange or red) topic, where the amount of color in each vertical bin presents the probability of the specific topic. The colors are comparable between diagrams. The manually scored hypnograms are provided below the topic diagrams. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

non-NDD cases, respectively. TN and FP represent the number of control subjects and PLM patients classified as non-NDD and NDD cases, respectively. The measures of the area under the receiver operation curve (AUC) were computed as the area under the curve obtained when going through the ranked probability measures. The regression model including four EEG features yielded estimates of 93.8% sensitivity and 87.5% specificity for the training dataset (filled circles in Figs. 7 and 8), and 91.4% sensitivity and 68.8% specificity for the validation dataset (open circles in Figs. 7 and 8). The regression model including two EOG features estimated a sensitivity of 65.6% and a specificity of 87.5% for the training dataset and a sensitivity of 57.1% and a specificity of 62.5% for the validation dataset.

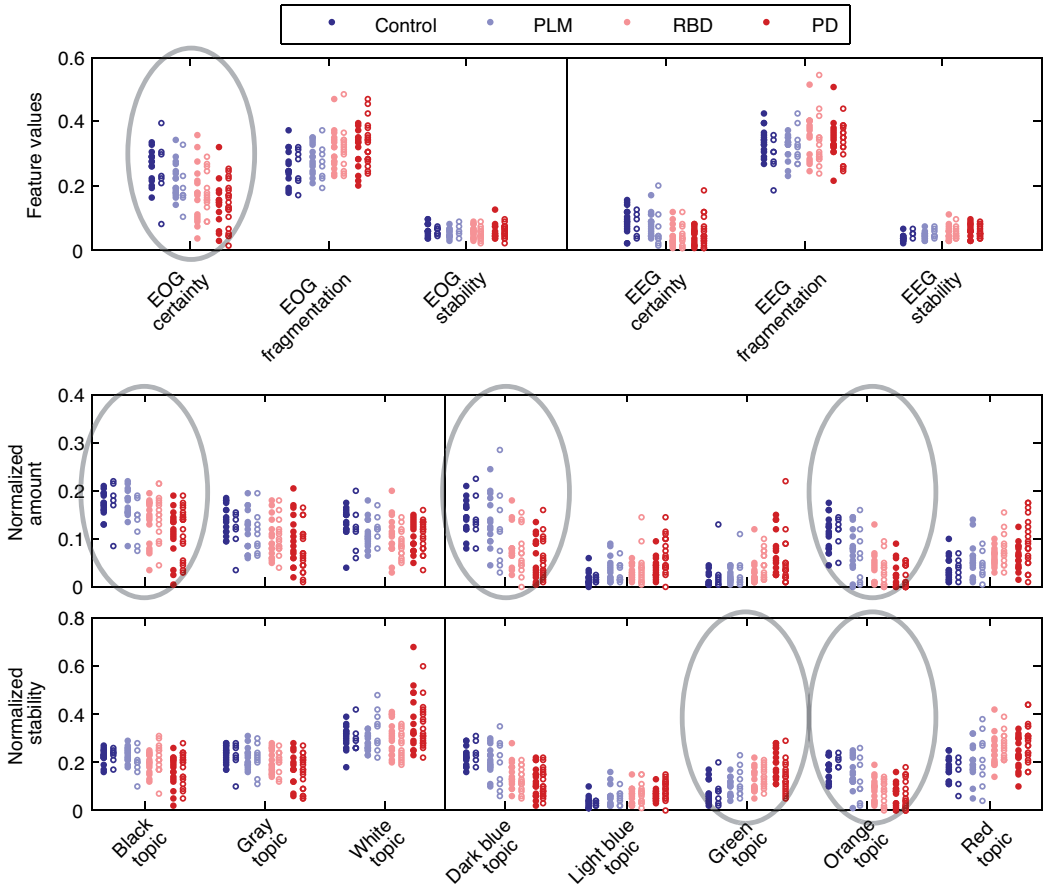


Fig. 7. The feature values obtained for the training and validation datasets. Subjects included in the validation and training datasets are presented as open and filled circles, respectively. Markers indicate control subjects (dark-blue), PLM patients (light-blue), iRBD patients (light-red) and PD patients (dark-red). The features chosen for the final regression models are indicated by gray, unfilled circles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 8 also illustrates the probabilities of belonging to the NDD class as given by the two regression models using standardized feature values. Subjects included in the training of the regression models are indicated by filled circles, and open circles indicate subjects in the validation dataset. The regression model initially including all features and that initially including solely EEG features were identical and so only one is shown. The results show that the distribution of the PLM patients in the validation set was relatively different from that of the training dataset. Together with the small amount of control and PLM patients in the validation set, this change in distributions is considered the main reason for the small specificity measures obtained in this study.

Table 5
AUC, sensitivity and specificity values for classifying PD and iRBD patients using the best logistic regression model obtained when (1) all features (EOG and EEG), (2) only EOG features or (3) only EEG features were available initially. The optimal subsets of features in each case are also presented. The same regression model was obtained for cases (1) and (3).

Input features		Feature subset	Performance measure	
EOG	EEG		Training dataset	Validation dataset
X	X	1: EEG topic "dark-blue" amount (REM amount) 2: EEG topic "orange" stability (N3 stability) 3: EEG topic "orange" amount (N3 amount) 4: EEG topic "green" stability (N1/N2 stability)	AUC: 93.1% Sensitivity: 93.8% Specificity: 87.5%	AUC: 84.3% Sensitivity: 91.4% Specificity: 68.8%
X		1: EOG topic "black" amount (N3 amount) 2. EOG overall certainty	AUC: 79.3% Sensitivity: 65.6% Specificity: 87.5%	AUC: 64.6% Sensitivity: 57.1% Specificity: 62.5%
	X	1: EEG topic "dark-blue" amount (REM amount) 2: EEG topic "orange" stability (N3 stability) 3: EEG topic "orange" amount (N3 amount) 4: EEG topic "green" stability (N1/N2 stability)	AUC: 93.1% Sensitivity: 93.8% Specificity: 87.5%	AUC: 84.3% Sensitivity: 91.4% Specificity: 68.8%

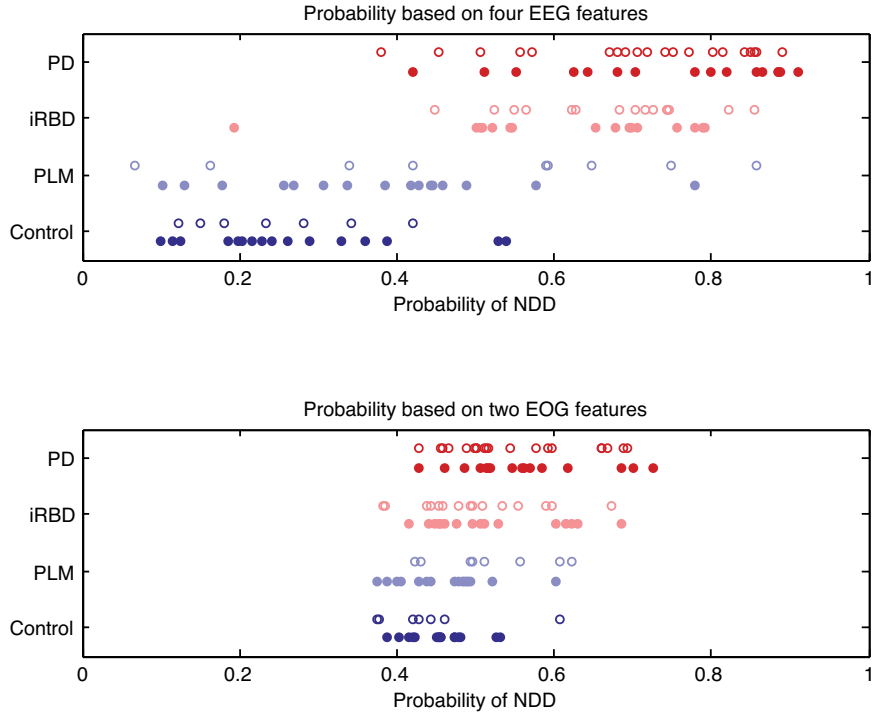


Fig. 8. Probabilities of being NDD as indicated by the best regression models obtained. The closer a subject is to the value 1 on the *x*-axis, the higher probability of belonging to the "NDD" class. The upper figure shows results from a model that includes four EEG features: $f_{T_{\text{Amount}}}^{\text{Dark Blue}}$ (amount of ~REM), $f_{T_{\text{Amount}}}^{\text{Orange}}$ (amount of ~N3), $f_{T_{\text{Stability}}}^{\text{Orange}}$ (stability of ~N3) and $f_{T_{\text{Stability}}}^{\text{Green}}$ (stability of ~N1/N2). The lower figure shows results from a model including two EOG features: $f_{\text{Certainty}}$ (certainty of any EOG topic) and $f_{T_{\text{Amount}}}^{\text{Black}}$ (amount of ~N3).

The non-NDD groups contained more females and slightly younger subjects than the NDD groups, and therefore the results were additionally validated by analyzing different subgroups of the data. Table 6 illustrates the performance measures obtained from the model including four EEG features for (1) subjects aged younger than 60 years, (2) subjects aged 60 years or more, (3) females only and (4) males only. It may be seen that the model performed best for the female subset and worst for the male subset.

4. Discussion

We propose a complete data-driven approach to reveal the immanent states in sleep EEG and EOG. Without any prior information such as manual sleep staging or subjective inputs, the immanent states are found and features reflecting EEG and EOG characteristics are derived. These characteristics are used to classify iRBD and PD patients automatically using a training dataset to build classifiers and a validation dataset to evaluate the classifiers' performance. Our study has three major findings. Firstly, it was found that features reflecting the amount and stability of two EEG topics similar to N3 and REM sleep, respectively, were specific to iRBD or PD. The feature reflecting amount of REM sleep was the more sensitive of the two, and the features reflecting the amount and stability of N3 sleep and the stability of a topic linked to N1/N2 were supportive. Secondly, it was found that by using features extracted from the EOG alone, two features, reflecting the amount of an EOG

Table 6
AUC, sensitivity and specificity values for classifying PD and iRBD patients in four subsets of the data. The best logistic regression model, which includes four EEG features, was used.

Dataset subgroup	Counts [total (control + PLM/iRBD + PD)]		Performance measures	
	Training dataset	Validation dataset	Training dataset	Validation dataset
Subjects aged <60 years	28 (21/7)	18 (9/9)	AUC: 99.3% Sensitivity: 100% Specificity: 95.2%	AUC: 87.7% Sensitivity: 88.9% Specificity: 77.8%
Subjects aged ≥60 years	36 (11/25)	33 (7/26)	AUC: 87.3% Sensitivity: 92.0% Specificity: 72.7%	AUC: 79.1% Sensitivity: 92.3% Specificity: 57.1%
Female subjects	26 (18/8)	19 (10/9)	AUC: 88.9% Sensitivity: 75.0% Specificity: 94.4%	AUC: 91.1% Sensitivity: 77.8% Specificity: 80.0%
Male subjects	38 (14/24)	32 (6/26)	AUC: 91.7% Sensitivity: 100% Specificity: 78.6%	AUC: 71.8% Sensitivity: 96.2% Specificity: 50.0%

topic related to N3 and the overall certainty of EOG, classified the patients with 57.1% sensitivity and 62.5% specificity, which are considered low performance measures. Thirdly, it was found that the features derived from the EEG topic models were better at classifying iRBD and PD patients than were the features derived from the EOG topic models. The best regression model included four EEG features and classified the iRBD and PD patients with 91.4% sensitivity and 68.8% specificity. The low specificity measures are thought to be caused by the skewed distribution of the PLM patients between the training and validation datasets. Overall, this study highlights the potential of the data-driven characterization of sleep EEG and EOG for evaluating neurodegenerative patients.

As may be seen in Fig. 7, the stability of the orange EEG topic (∼N3 sleep) and the dark-blue EEG topic (∼REM sleep) were lower in iRBD and PD patients. Controls and PLM patients had greater stability, expressed as the higher frequency of successive sleep epochs with a dominant probability of N3 or REM. These findings suggest that the sleep-regulating system in iRBD and PD patients, which may involve REM–NREM transition (Lu et al., 2006), is affected. Loss of atonia during REM sleep is a diagnostic criterion for RBD, and the neurons regulating atonia are probably affected in these patients. REM sleep in this study was identified on the basis of EEG "words", and the finding that iRBD patients have lower REM stability suggests that the neurons regulating EEG components of REM sleep are also affected.

The amount of the dark-blue EEG topic (∼REM) and the orange EEG topic (∼N3) are both discriminative features, as shown in Fig. 7. It should be emphasized that the features reflect the amount of "certain" orange and "certain" dark-blue topic, as only sleep epochs with a probability of more than 70% of the given topic are included in the feature value. NDD patients have lower probabilities of these topics in many sleep epochs, which may be explained by the EEG word combination. The lower probabilities of orange topics in these patients could be caused by greater sleep fragmentation, and they thus rarely display a period of maintained delta waves, as do control subjects during N3 sleep. Additionally, the deep structures and/or the integrity of the brain could be affected in such a way that patients simply cannot generate and/or maintain delta wave sleep. Presuming that a delta wave is represented as a given word used in the EEG topic model, the absence or low frequency of delta waves in an epoch will be reflected by a lower probability of topic orange.

Another reason for the lower probability of the orange and dark-blue topics in iRBD and PD patients could be that the epochs are contaminated with different mixtures of microsleep events. Each microsleep event is thought to be presented as a given word and in the process of training the topic models, a certain distribution of words is acquired from control subjects. If the distributions of microsleep events in patients differ significantly from the distributions in controls, the topic models are unable to recognize the distribution over words (the "fingerprint" of the epoch), which will be reflected as lower probabilities of the topics.

A third explanation of the lower probabilities seen in the patients could be that microsleep events in patients are changed or affected to such a degree that they are not expressed as the same words as in control subjects. The microsleep events themselves as well as the distribution of them form specific fingerprints of the sleep stages. If the "vocabulary" is not the same for patients and control subjects, the topic models might not recognize the word distribution in patients and thereby provide lower probabilities of the topics. Whether the reduced amount of the dark-blue and orange topic is caused by changes in the individual words or in the distribution of words has to be investigated thoroughly in future studies. Such a study could also investigate the influence of topography of EEG words on topic probabilities.

The most indicative EOG feature was found to be the amount of the black topic, which is believed to be related to the low-frequency high-amplitude EEG artifacts present in the EOG signals during N3 sleep. This result is consistent with the findings of the most indicative EEG features, as these also showed that the amount of N3 sleep was indicative of PD/iRBD. It is suggested that the black topic (EOG) reflects the same characteristic as the orange topic (EEG), as both are pronounced during N3 sleep. This may explain why no EOG features were found to be supportive of the EEG features. The EOG topic diagrams do not show as complex and varied a sleep structure as the EEG topic diagrams. Therefore, it is concluded that even though the overall sleep pattern can be described by the use of EOG alone, potential EOG changes in PD/iRBD patients are not captured or identified in this study. However, the feature reflecting overall EOG certainty was supportive in the classification model, indicating that iRBD and PD patients do not fall into the normal states found for EMs.

Future studies should investigate whether the findings of this study are specific to iRBD and PD patients. If the reduced amount of the orange EEG topic (∼N3) in patients is caused by lighter sleep and a lower frequency of delta waves than in control subjects, this indicative feature might not be specific to iRBD and PD. Additionally, lighter and more abrupted sleep could also explain the slightly greater amounts of the dark-blue EEG topic (∼REM) in control and PLM patients compared with iRBD and PD patients. If this is the case, other patients suffering from apnea, or control subjects sleeping poorly, might also be classified as "patients" if these features alone were used in the classification scheme. Additionally, gender and age affect sleep efficiency and sleep architecture, including the amount and stability of NREM and especially N3 sleep. Although the groups in this study were defined to be matched by age in the training dataset, there are still dissimilarities between the control/PLM group and the iRBD/PD group, which may enhance the observed differences. Table 6 provides performance results for the model, which includes four EEG features when evaluating (1) subjects aged less than 60 years, (2) subjects aged 60 years or more, (3) only females and (4) only males. Some of the data subgroups were not well matched between the two groups and this should be kept in mind when interpreting the results. Although the model seems to perform best in younger female subjects, the AUC measures of the validation data indicate that the findings in this study were not solely due to age and gender effects.

This study did not investigate which words or word combinations were specific to the different topics. Neither did we investigate how much nor in which way the different topics differed from the AASM stages. An advantage of evaluating sleep using this data-driven method is that it is fully automatic and does not use any prior information. The method mimics manual scoring as it encompasses words reflecting microsleep patterns across different topographic sites. However, it differs from manual scoring because each epoch is independent of adjacent epochs and presented as a mixture of topics rather than discrete stages. The features presented are to a certain degree dependent on the topic probabilities, but additional features could be developed that take even more advantage of the topic mixtures and topic probabilities. For instance, the thresholds involved in the certainty features and the topic amount features could be investigated in depth and the concurrency of the different topics determined. Finally, future studies could include an analysis of how well the features perform if they are not computed on the basis of topics, but simply from the manually scored hypnogram.

The data-driven approach presented here overcomes the evaluation of the macrostructure of sleep and gives the opportunity to find words and topics that are not obvious to the human eye. Although the EOG features did not discriminate the iRBD/PD and control/PLM groups, we believe that a data-driven analysis of

EOG could provide helpful information for evaluating sleep and sleep-related disorders. A suggested future study may include a refinement of the EOG topic models and a thorough evaluation of what the various EOG topics reflect. Ultimately, if the three EOG topics reflect proportions of SEMs, REMs and no EMs, the temporal distribution and mixture of slow/fast/none EMs could be investigated. We believe that the features presented here derived from such EOG topic diagrams could be indicative of neurodegeneration in iRBD and PD patients. Additionally, the features could be helpful when evaluating other patients, as EMs may be affected differently in distinct NDDs. In Alzheimer's disease (AD), for instance, the loss of neurons and synapses occurs mainly in the cortex and certain subcortical regions (Wenk, 2003). Given this, it can be hypothesized that saccade initiation (reflected in the REMs), which is mainly controlled by neurons located in the frontal brain areas, is more affected in AD patients than in pre-PD patients. We may also speculate that the autonomous phase of saccades (reflected in the SEMs), which is mainly controlled by neurons located in the brain stem, is more affected in pre-PD patients than in AD patients. In this way, different types of EMs are thought to be affected in different NDDs, and as these differences could be reflected in the EOG word combination, the EOG features derived in this study could be helpful for refining the specificity of different NDDs. However, this needs to be confirmed in future studies.

Finally, it should be emphasized that the frequency of high probabilities of either topic is reflected in the "certainty" measures. It may be seen in Fig. 7 that control and PLM patients display slightly higher values of the "EEG certainty" and "EOG certainty" features than do the iRBD and PD patients. These features can be seen as a reflection of how well each test subject fits into the "control" topic models, and they conclusively demonstrate that the iRBD and PD patients reflect abnormal form and/or time distribution of both EMs and EEG structures during sleep. We believe that our findings could support the well-known lack of atonia, which is reflected in the EMG of these patients. Combining a variety of PSG biomarkers would enhance the sensitivity and specificity when automatically detecting subjects at risk of developing PD and other NDDs. Additionally, combining several specific PSG findings would give a more detailed description of the individual subjects and could serve as a basis for specific patient medication and potentially efficacious treatment regimes.

In a related study (Koch et al., 2014), we took the same topic modeling approach using the same PSG data, but for a different purpose. We focused on optimizing the topic model through supervised learning to identify and describe the standard sleep stages through the use of EEG and EOG. In combination, the two studies indicate that a data-driven topic modeling approach can be useful for analyzing sleep – either by finding characteristics indicative of a certain disorder, or by automatically identifying and elaborately describing the standard sleep stages using only EEG and EOG.

5. Conclusion

We propose a data-driven approach for characterizing sleep EEG and EOG, where topics are identified unsupervised by learning structures directly from the data. Two topic models were built based on EEG or EOG data, both of which enabled the standard sleep stages to be visually identified. By extracting and using features reflecting characteristics of the topics found, unseen (validation) patients with a diagnosis of iRBD or PD were classified with a sensitivity of 91.4% and a specificity of 68.8%. The features we developed indicate that sleep characteristics computed from sleep topics identified by a data-driven approach are indicative of neurodegeneration. Specifically, the normalized amount and stability of topics linked to N3 and REM were found to be considerably lower for iRBD and PD patients. Furthermore, the study suggests that the stability of a topic linked to N1/N2 is supportive of the classification scheme. We conclude that the amount of REM and N3, as well as the ability to maintain NREM and REM sleep determined by data-driven topic models can be used as potential early PD biomarkers.

Conflict of interest

The Ph.D. project is supported by grants from H. Lundbeck A/S, the Technical University of Denmark and the Center for Healthy Ageing, University of Copenhagen, Denmark. Julie A.E. Christensen, Marielle Zoetmulder, Henriette Koch, Rune Frandsen, Lars Arvastson, Søren R. Christensen, Poul Jennum and Helge B.D. Sorensen report no conflicts of interest.

References

Agarwal R, Takeuchi T, Laroche S, Gotman J. Detection of rapid-eye movements in sleep studies. IEEE Trans Biomed Eng 2005;52(8):1390–6.
Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. J Mach Learn Res 2003;3:993–1022.
Brown RE, Basheer R, McKenna JT, Strecker RE, McCarley RW. Control of sleep and wakefulness. Physiol Rev 2012;92(3):1087–187.
Braak H, Del Tredici K, Rüb U, de Vos RA, Steur EN, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging 2003;24(2):197–211.
Carpenter RH. The neural control of looking. Curr Biol 2000;10(8):R291–3.
Christensen JAE, Frandsen R, Kempfner J, Arvastson L, Christensen SR, Jennum P, et al. Separation of Parkinson's patients in early and mature stages and control subjects using one EOG channel. In: Conf Proc IEEE Eng Med Biol Soc 2012; 2012. p. 2941–4.
Christensen JAE, Koch H, Frandsen R, Kempfner J, Arvastson L, Christensen SR, et al. Classification of iRBD and Parkinson's disease patients based on eye movements during sleep. In: Conf Proc IEEE Eng Med Biol Soc 2013; 2013. p. 441–4.
Christensen JAE, Kempfner J, Zoetmulder M, Arvastson L, Christensen SR, Sorensen HBD, et al. Decreased sleep spindle density in patients with idiopathic REM sleep behaviour disorder and patients with Parkinson's disease. Clin Neurophysiol 2014;125(3):512–9.
Corin MS, Elizan TS, Bender MB. Oculomotor function in patients with Parkinson's disease. J Neurol Sci 1972;15(3):251–65.
Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. 2nd ed. New York: Springer; 2008.
Iber C, Ancoli-Israel S, Chesson AL, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events: rules, terminology, and technical specification. Am Acad Sleep Med 2007.
Iranzo A, Molinuevo JL, Santamaría J, Serradell M, Martí MJ, Valldeoriola F, et al. Rapid-eye-movement sleep behaviour disorder as an early marker for a neurodegenerative disorder: a descriptive study. Lancet Neurol 2006;5(7):572–7.
Kempfner J, Sorensen G, Zoetmulder M, Jennum P, Sorensen HBD. REM behaviour disorder detection associated with neurodegenerative diseases. In: Conf Proc IEEE Eng Med Biol Soc 2010; 2010. p. 5093–6.
Koch H, Christensen JAE, Arvastson L, Christensen SR, Jennum P, Sorensen HBD. Classification of iRBD and Parkinson's patients using a general data-driven sleep staging model build on EEG. In: Conf Proc IEEE Eng Med Biol Soc 2013; 2013. p. 4275–8.
Koch H, Christensen JAE, Frandsen R, Arvastson L, Christensen SR, Jennum PJ, et al. Automatic sleep classification using a data-driven topic model reveals latent sleep states. J Neurosci Methods 2014., http://dx.doi.org/10.1016/j.jneumeth.2014.07.002.
Latreille V, Carrier J, Montplaisir J, Lafortune M, Gagnon J-F. Non-rapid eye movement sleep characteristics in idiopathic REM sleep behavior disorder. J Neurol Sci 2011;310(1–2):159–62.
Lu J, Sherman D, Devor M, Saper CB. A putative flip-flop switch for control of REM sleep. Nature 2006;441(7093):589–94.
Mosimann UP, Müri RM, Burn DJ, Felblinger J, O'Brien JT, McKeith IG. Saccadic eye movement changes in Parkinson's disease dementia and dementia with Lewy bodies. Brain 2005;128(Pt 6):76–1267.
Postuma RB, Gagnon JF, Montplaisir J. Clinical prediction of Parkinson's disease: planning for the age of neuroprotection. J Neurol Neurosurg Psychiatry 2010a;81(9):1008–13.
Postuma RB, Gagnon JF, Rompré S, Montplaisir JY. Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease. Neurology 2010b;74(3):239–44.
Rodrigues Brazète J, Montplaisir J, Petit D, Postuma RB, Bertrand J-A, Génier Marchand D, et al. Electroencephalogram slowing in rapid eye movement sleep behavior disorder is associated with mild cognitive impairment. Sleep Med 2013;14(11):1059–63.
Saper CB, Chou TC, Scammell TE. The sleep switch: hypothalamic control of sleep and wakefulness. Trends Neurosci 2001;24(12):726–31.

Saper CB, Fuller PM, Pedersen NP, Lu J, Scammell TE. Sleep state switching. Neuron 2010;68(6):1023–42.

Schenck CH, Boeve BF, Mahowald MW. Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series. Sleep Med 2013;14(8):744–8.

Schenck CH, Bundlie SR, Mahowald MW. Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder. Neurology 1996;46:388–93.

Schwartz JRL, Roth T. Neurophysiology of sleep and wakefulness: basic science and clinical implications. Curr Neuropharmacol 2008;6(4):367–78.

Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc Series B Stat Methodol 1996;58(1):267–88.

Van Esbroeck A, Westover B. Data-driven modeling of sleep states from EEG. In: Conf Proc IEEE Eng Med Biol Soc 2012; 2012. p. 5090–3.

Verbeek J. Matlab Latent Dirichlet Allocation toolbox; 2006 http://lear.inrialpes.fr/~verbeek/software.php

Wenk GL. Neuropathologic changes in Alzheimer's disease. J Clin Psychiatry 2003;64(9):7–10.

Paper IV

TITLE:
Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease

AUTHORS:
Julie A. E. Christensen, Poul Jennum, Henriette Koch, Rune Frandsen, Marielle Zoetmulder, Lars Arvastson, Søren R. Christensen and Helge B. D. Sorensen

JOURNAL:
Clinical Neurophysiology

YEAR:
2015

VOLUME:
[in press]

PAGES:
[in press]

PUBLICATION HISTORY:
Submission date: 23 September 2014, acceptance date: 10 March 2015

Clinical Neurophysiology xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease ☆

Julie Anja Engelhard Christensen [a,b,c,*], Poul Jennum [b], Henriette Koch [a], Rune Frandsen [b], Marielle Zoetmulder [b,d], Lars Arvastson [c], Søren Rahn Christensen [c], Helge Bjarrup Dissing Sorensen [a]

[a] Department of Electrical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark
[b] Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, University of Copenhagen, Glostrup Hospital, Glostrup, Denmark
[c] H. Lundbeck A/S, Copenhagen, Denmark
[d] Department of Neurology, Bispebjerg Hospital, Copenhagen, Denmark

ARTICLE INFO

Article history:
Accepted 10 March 2015
Available online xxxx

Keywords:
Parkinson's disease
Rapid eye movement sleep behavior disorder
Polysomnography
Sleep transitions
Data-driven detection

HIGHLIGHTS

- Wake and sleep stability is affected in early stages of Parkinson's disease.
- An automated REM stability index is a potential biomarker for Parkinson's disease.
- We discuss problems linked to micro-sleep and nocturnal sleep fragmentation.

ABSTRACT

Objective: Patients with idiopathic rapid eye movement (REM) sleep behavior disorder (iRBD) are at high risk of developing Parkinson's disease (PD). As wake/sleep-regulation is thought to involve neurons located in the brainstem and hypothalamic areas, we hypothesize that the neurodegeneration in iRBD/PD is likely to affect wake/sleep and REM/non-REM (NREM) sleep transitions.
Methods: We determined the frequency of wake/sleep and REM/NREM sleep transitions and the stability of wake (W), REM and NREM sleep as measured by polysomnography (PSG) in 27 patients with PD, 23 patients with iRBD, 25 patients with periodic leg movement disorder (PLMD) and 23 controls. Measures were computed based on manual scorings and data-driven labeled sleep staging.
Results: Patients with PD showed significantly lower REM stability than controls and patients with PLMD. Patients with iRBD had significantly lower REM stability compared with controls. Patients with PD and RBD showed significantly lower NREM stability and significantly more REM/NREM transitions than controls.
Conclusions: We conclude that W, NREM and REM stability and transitions are progressively affected in iRBD and PD, probably reflecting the successive involvement of brain stem areas from early on in the disease.
Significance: Sleep stability and transitions determined by a data-driven approach could support the evaluation of iRBD and PD patients.

1. Introduction

Parkinson's disease (PD) and other synucleinopathies are debilitating diseases with impacts on morbidity, mortality, work, and social and family life. The disorders have large direct and indirect costs for society (Jennum et al., 2011). PD is the second most common neurodegenerative disease after Alzheimer's disease. Treatment is purely symptomatic and does not alter underlying disease progression (Schapira et al., 2013). When motor symptoms are present, alterations of the substantia nigra with reduced production or depletion of dopamine is found (Galvin et al., 2001),

Abbreviations: AASM, American Academy of Sleep Medicine; DCSM, Danish Center for Sleep Medicine; EOG, electrooculography; iRBD, Idiopathic rapid eye movement sleep behavior disorder; MSLT, Multiple sleep latency test; NREM, Non-rapid eye movement sleep; PD, Parkinson's disease; PLMD, Periodic leg movement disorder; PSG, Polysomnography; RBD, Rapid eye movement sleep behavior disorder; REM, Rapid eye movement sleep; RSWA, REM sleep without atonia; SLD, Sublaterodorsal tegmental nucleus; SS, Sleep spindles; W, Wake.

☆ Institution at which the work was performed: Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Glostrup Hospital, Glostrup, Denmark.

* Corresponding author at: Technical University of Denmark (work), Orsteds Plads, Building 349, 2800 Kongens Lyngby, Denmark. Tel.: +45 45255737 (work), Mobile: +1 650 771 8073 (private).
E-mail address: julie.a.e.christensen@gmail.com (J.A.E. Christensen).

http://dx.doi.org/10.1016/j.clinph.2015.03.006
1388-2457/

Please cite this article in press as: Christensen JAE et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.03.006

mediating the classical PD phenotype. Over the last 20 years, it has become increasingly clear that the entire brain is affected by the pathology, which typically starts in caudal areas of the brainstem and progresses anteriorly (Braak et al., 2003).

Recent research has focused on early detection of PD, notably changes in sleep–wake pattern and rapid eye movement (REM) sleep without atonia (RSWA) coinciding with dream-enacting behavior – REM sleep behavior disorder (RBD) (Salawu et al., 2010). RBD is closely associated with PD (Munhoz and Teive, 2014), and patients suffering from idiopathic RBD (iRBD) are at great risk of subsequently developing Parkinsonism or dementia (Schenck et al., 2013, 2003, 1996). Several studies have focused on analysis of sleep data in the search for PD biomarkers (Dos Santos et al., 2014). These have examined measures of sleep spindle densities, RSWA, slow wave characteristics (Christensen et al., 2014b; Kempfner et al., 2014a,b; Latreille et al., 2014, 2011; Postuma et al., 2010) and other measures of abnormalities of brain stem function, including autonomic functions such as heart-rate variability (Sorensen et al., 2013a, 2012) and other non-motor symptoms (Garcia-Ruiz et al., 2014; Sakakibara et al., 2014).

Sleep is strongly regulated by groups of neurons located in the brainstem and midbrain areas, which form reciprocal connections (Luppi et al., 2011; Saper et al., 2010, 2001; Schwartz and Roth, 2008). These "sleep–wake switches" are mutually dependent and have been referred to as the wake–sleep and REM–NREM sleep switches, respectively. Despite the mutually inhibitory loops involved in the two switching mechanisms, if either side of the two loops is weakened or injured, unwanted instability can occur in either of the states, irrespective of which side is damaged (Schwartz and Roth, 2008). As neurons of the brain stem and basal brain structures are affected in synucleinopathies (Braak et al., 2003), we propose that the neurodegeneration will have a progressive impact on the wake–sleep and REM–NREM transitions and stability. Manual scoring of sleep in patients with PD is not very reliable and prone to high inter- and intra-rater variability (Danker-Hopfe et al., 2004; Jensen et al., 2010). Therefore, this study analyzed wake–sleep and REM–NREM transitions as well as W, REM and NREM stability measures based on automatically identified as well as manually scored REM, NREM and W stages. The automatic method used has been validated by (Koch et al., 2014) using the same PSG data as analyzed in this study.

2. Methods

2.1. Subjects and recordings

Subjects were recruited from the Danish Center for Sleep Medicine (DCSM) in the Department of Clinical Neurophysiology, Glostrup University Hospital in Denmark. A total of 27 patients with PD, 23 patients with iRBD, 25 patients with periodic leg movement disorder (PLMD) and 23 control subjects aged 40 years or more and with no history of movement disorder, dream-enacting behavior or other previously diagnosed sleep disorders were included. Nineteen of the patients with PD had RBD (PD^+) and eight did not (PD^-), as determined by the presence of RSWA as well as clinical complaints. All or a subset of the subjects included here have appeared in previous studies (Christensen et al., 2014a,b; Sorensen et al., 2013a; Zoetmulder et al., 2014a,b). Patient evaluations included a comprehensive medical and medication history and a polysomnography (PSG) analyzed according to the American Academy of Sleep Medicine (AASM) standard (Iber et al., 2007). A multiple sleep latency test (MSLT) was performed in any cases where narcolepsy was suspected. The PD diagnose rely on clinical features including motor information typically for PD which further includes DAT scan and in some cases also MRI of the brain. The RBD Screening Questionnaire (RBDSQ) (Stiasny-Kolster et al., 2007) was used to screen for RBD, and the iRBD patients were divided in two groups: (1) those with a total score of nine or less ($iRBD^-$) and (2) those with a total score of 10 or more ($iRBD^+$). The cutoff was chosen to divide the iRBD patients into those with major self-reported dream enactments and those with minor. All iRBD patients thus had self-reported dream enactment as well as RSWA in the recorded night analyzed. Patients treated with medication known to affect sleep stages (antidepressants, antipsychotics, hypnotics) were excluded, except for dopaminergic treatments. We are fully aware of the potential effect of dopaminergic drugs on vigilance (Micallef et al., 2009), but discontinuation of dopaminergic treatment in actively treated patients prior to PSG also risks deleterious discontinuation effects. Furthermore, discontinuation is very difficult to achieve in clinical settings and may have unpleasant and negative motor effects that could interfere with the study and might even be unethical. One patient in the PD group later developed Multiple System Atrophy and another patient with PD developed Lewy Body Dementia. Controls underwent at least one night of PSG as an outpatient. Patients underwent at least one night of PSG as an outpatient or in hospital in accordance with the AASM standard (Iber et al., 2007). The quality of each PSG data set was individually evaluated, and recordings were discounted if channels became disconnected or were significantly contaminated with artifacts. Low-quality PSGs were repeated whenever possible. Demographic data as well as PSG variables for the six groups are summarized in Table 1.

2.2. Automatic staging of sleep

Sleep changes in both microarchitecture and macroarchitecture have been reported in neurodegenerative disease, the former interfering with sleep scoring and increasing intra- and inter-rater variability. Specifically, patients with iRBD and PD have been reported to have EEG frequency slowing (Rodrigues Brazète et al., 2013), changed EEG during REM sleep (Christensen et al., 2014a; Fantini et al., 2003; Hansen et al., 2013), changed morphology or fewer rapid or slow eye movements (Christensen et al., 2014a, 2013) and sleep spindles (SS) (Christensen et al., 2014b; Latreille et al., 2014), and display RSWA. The listed findings contribute to very altered sleep and consequently, sleep stage scoring this pathology is associated with high inter-rater variability (Danker-Hopfe et al., 2004; Jensen et al., 2010). To overcome this, an automatic sleep detector was used to identify REM, NREM and W for each subject and analyses were performed using automatically scored sleep stage data. The automatic sleep scoring technique used in this study has been validated on the same PD, iRBD, PLMD and control PSG data set as used in this study, and the methods are described in detail by Koch et al. (2014). Specifically, the method is optimized on nocturnal PSG of 50 subjects, and validated on an additional 76 subjects (a mixture of the same controls and patients included in this study). Our automatic detector is data-driven, identifying sleep states based on the distribution of certain EEG and electrooculographic (EOG) characteristics, which can be interpreted as specific fingerprints for each epoch. Due to the presence of RSWA, the automatic identification approach is based on EEG and EOG characteristics only. Briefly, for each second, EEG C3-A2 and O1-A2 single-sided amplitude spectra in clinical frequency bands as well as EOG power below 5 Hz and cross-correlations between the two EOG channels are computed. Surveying these measures simultaneously in three-second intervals with a step size of one second, the approach identifies patterns indicative of the various sleep stages. The method produces a mixture of probabilities for the different sleep stages, and final identification of NREM, REM and W is based on the highest probability when combining probabilities of individual sleep stages (Koch et al., 2014). A clear

Please cite this article in press as: Christensen JAE et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.03.006

Table 1
Demographic and PSG data for the six groups studied. Patients with Parkinson's disease (PD) were divided in those with REM sleep behavior disorder (RBD) (PD^+) and those without (PD^-), as determined by the presence of REM sleep without atonia as well as clinical complaints. The RBD Screening Questionnaire was used to divide the patients with idiopathic RBD (iRBD) in those with a total score of ⩽9 ($iRBD^-$) and those with a total score of >10 ($iRBD^+$). The patients with periodic leg movement disorder (PLMD) were included as a secondary control group. The disease onset is stated as years from clinical diagnosis (PD patients) or self-reported subjective RBD-symptoms (iRBD patients).

	Controls	PLMD	$iRBD^-$	$iRBD^+$	PD^-	PD^+
Total counts (Male/Female)	23 (7/16)	25 (13/12)	12 (9/3)	11 (10/1)	8 (5/3)	19 (13/6)
Age [years, $\mu \pm \sigma$]	56.7 ± 9.2	56.9 ± 11.6	61.8 ± 6.8	66.3 ± 7.2	68.8 ± 8.4	63.7 ± 6.7
BMI [kg/m^2, $\mu \pm \sigma$]	23.1 ± 2.5	25.7 ± 3.8	25.1 ± 2.6	25.9 ± 3.6	24.1 ± 3.6	25.9 ± 2.9
Sleep Efficiency [%, $\mu \pm \sigma$]	86.2 ± 10.9	86.1 ± 7.5	82.5 ± 7.4	85.8 ± 6.7	67.9 ± 15.5	80.3 ± 9.9
Time in Bed [min, $\mu \pm \sigma$]	484 ± 81.2	431 ± 50.0	518 ± 107.3	474 ± 71.1	427 ± 53.8	461 ± 88.3
W [%, $\mu \pm \sigma$]	13.8 ± 10.9	13.9 ± 7.5	17.5 ± 7.4	14.2 ± 6.7	32.1 ± 15.5	19.7 ± 9.9
REM [%, $\mu \pm \sigma$]	18.9 ± 6.6	18.5 ± 6.1	17.7 ± 7.1	16.3 ± 7.1	9.3 ± 5.3	14.6 ± 10.8
N1 [%, $\mu \pm \sigma$]	8.8 ± 4.6	8.4 ± 7.2	7.5 ± 2.6	10.8 ± 5.4	7.3 ± 6.0	12.6 ± 9.9
N2 [%, $\mu \pm \sigma$]	44.7 ± 9.4	44.7 ± 10.9	40.0 ± 10.9	43.0 ± 10.3	39.2 ± 12.9	39.3 ± 15.1
N3 [%, $\mu \pm \sigma$]	13.8 ± 7.3	14.5 ± 9.1	17.3 ± 9.9	15.7 ± 11.2	12.1 ± 9.1	13.9 ± 17.6
LM index [no/hour, $\mu \pm \sigma$]	18.1 ± 13.1	56.5 ± 36.8	49.9 ± 32.0	31.6 ± 22.9	39.2 ± 40.1	42.7 ± 44.4
PLM index [no/hour, $\mu \pm \sigma$]	7.7 ± 7.4	36.5 ± 24.6	14.9 ± 16.3	23.0 ± 17.1	11.2 ± 11.5	10.8 ± 10.3
Disease duration [years, $\mu \pm \sigma$]	NA	NA	5.3 ± 9.8	11.3 ± 12.0	7.7 ± 5.7	4.6 ± 3.3
RBD score [$\mu \pm \sigma$]	NA	NA	7.7 ± 1.6	11.1 ± 1.0	2.5 ± 1.1	9.7 ± 2.3

advantage of this approach is that it takes distributions of micro-sleep characteristics into account, and each sleep epoch is labeled individually and independently of adjacent epochs. In spite of the pace of the sleep-regulating loops, the sleep transitions happen gradually, and the method used will assign an epoch a stage once the probability of this stage becomes the dominant one, and not when only few distinct signs of a stage is shown. It is also one of the reasons why the manually and automatically scored stages differ. The overall mean accuracy rates for detecting NREM, REM and W ranged from 70% for patients with PD to 77% for control subjects when comparing with manually single-scored hypnograms.

2.3. Analysis of transitions and stability

Two transition measures and three stability measures were defined and analyzed in this study. Wake–sleep transitions were defined as the number of shifts from any sleep stage (N1, N2, N3, REM) to wakefulness or *vice versa*. REM–NREM transitions were defined as the number of shifts from REM sleep to any NREM sleep stage (N1, N2, N3) or *vice versa*. Both measures were defined as the frequency per minute of total time in bed. The three stability measures were defined as the number of passages/toggles 1) from a REM stage to a REM stage (REM → REM), 2) from any NREM stage to any NREM stage (NREM → NREM) or 3) from a W stage to a W stage (W → W). These measures were defined as the frequency of passages between two REM, NREM or W epochs per minute of the total time spent in these stages, respectfully. For each subject, transition and stability measures were computed based on the manually scored hypnogram as well as the automatic REM/NREM/W staging technique. Wilcoxon rank-sum tests were performed to compare the between-group transition and stability measures, yielding 15 comparisons for each measure summing up to 75 tests in total. A significance level of $p < 0.05$ was used. The Benjamini–Hochberg procedure was used to correct for multiple testing using a false discovery rate at level $q = 0.10$.

3. Results

Transition and stability measures computed from the automatic identified NREM, REM and W are illustrated in Fig. 1, and the results found from the manually scored hypnogram are illustrated in Fig. 2. Mean and standard deviations are illustrated in bar plots to the left and measures for each subject illustrated as dots to the right.

The results obtained from the automatically identified stages showed that patients with iRBD with a RBDSQ total score of nine or lower had significantly lower REM stability compared with controls ($p = 0.0106$). Patients with iRBD with a RBDSQ total score of ten or more had significantly lower REM stability compared with controls ($p = 0.0036$) and PLMD patients ($p = 0.0091$). PD patients without RBD showed significantly lower REM stability compared to the controls ($p < 0.0055$) and PLMD ($p < 0.0052$). PD patients with RBD had significantly lower REM stability than controls ($p = 0.0003$) and PLMD patients ($p = 0.0004$), significantly lower NREM stability than controls ($p = 0.0050$), and significantly more REM–NREM sleep transitions than controls ($p = 0.0107$). Finally, iRBD patients with a RBDSQ total score of nine or lower showed a trend towards lower REM stability compared with PLMD patients, and iRBD patients with a RBDSQ total score of ten or more showed a trend of more REM–NREM sleep transitions than controls. However, these trends were not statistically significant after correction for multiple testing. No significant differences were found between any of the iRBD and PD groups, although PD patients showed a trend towards lower REM stability and more REM–NREM sleep transitions compared with iRBD patients.

No significant between-group differences were found for the transition and stability measures computed from the manually scored hypnograms. It is seen that the stability measures computed from the manually scored hypnogram are all greater than the measures computed from the data-driven sleep staging technique. The measures for REM–NREM sleep transitions are greater for the automatic staging technique compared to the manually

Please cite this article in press as: Christensen JAE et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.03.006

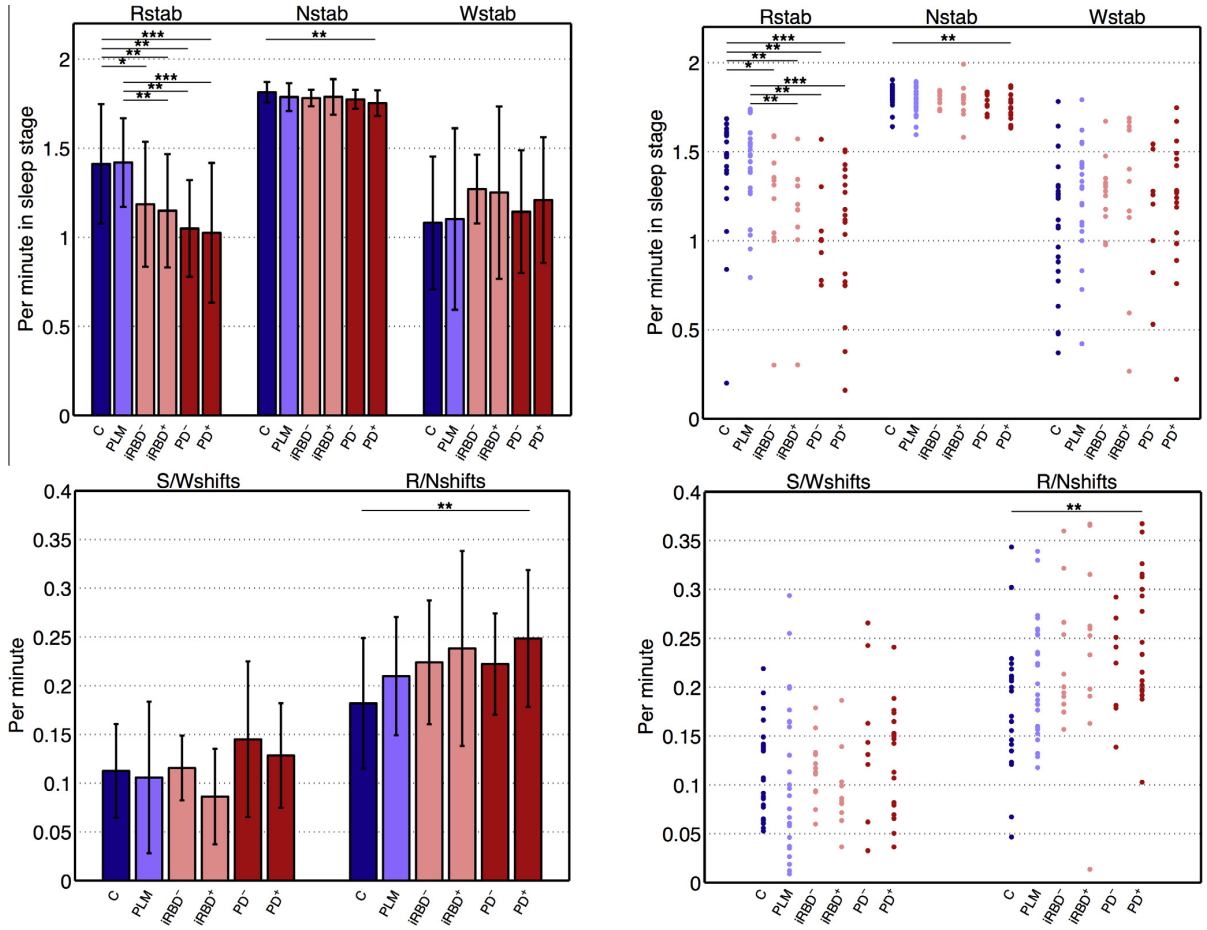


Fig. 1. The frequencies of wake–sleep and REM–NREM shifts and the stability measures for REM sleep, NREM sleep and wake for the six groups based on the automatically scored REM, NREM and W epochs. Left, bar charts with heights indicating means, and error bars indicating one standard deviation of the observations above and below the mean. Right, results from each subject indicated as dots. Dark blue, control subjects; light blue, PLMD patients; light red, iRBD patients; dark red, patients with Parkinson's disease (PD). Asterisks indicate a significant between-group difference, determined by Wilcoxon rank-sum tests: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. The Benjamini–Hochberg procedure was used to control the false discovery rate (FDR) at level $q = 0.10$. Fifteen between-group comparisons were made of five measures, giving a total of 75 tests. Only the results for the comparisons that remained significant after FDR correction are presented. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

scored sleep stages, and the differences between the REM stability measures are thought to be the main cause hereof.

4. Discussion

This study is the first to analyze wake–sleep and REM–NREM transitions and the stability of REM, NREM and W in PD, iRBD and PLMD patients compared with controls. Furthermore, we divided the PD patients into those with and those without RBD, and the iRBD patients were divided in those with a high total score on the RBDSQ and those with a low score. Our main findings are: (1) REM sleep is less stable in iRBD and PD patients than in PLMD patients and control subjects, regardless of the RBDSQ total score and the presence of RBD. A non-significant trend was seen for a lower REM sleep stability in PD compared to iRBD patients. (2) PD patients with RBD showed significantly lower NREM stability and significantly more REM–NREM shifts than controls. Same trend, however non-significant, was seen for iRBD with high scores in the RBDSQ compared with control subjects. Overall, trends were seen for lower REM stability and more REM/NREM transitions in both groups of iRBD and PD patients. These results indicate that the ability to maintain REM and NREM sleep is a biomarker for iRBD and PD, regardless of the presence of clinical symptoms.

The sub-division of iRBD and PD patients might not be optimal, as it does not reflect the actual RBD severity in the recording analyzed. The PD patients without RBD might have RSWA to a certain degree but no RBD diagnosis as they lack to report subjective symptoms. Contradictory, the iRBD patients with major self-reported dream enactments might show just sufficient RSWA to be diagnosed. Analyzing trends for the stability and transitional measures as a function of clinical RBD severity could give a better insight in how correlated the features are with RBD severity. The clinical RBD severity scale reported in (Sixel-Döring et al., 2011) analyzes RBD on an event-to-event basis, including both vocalization and movements, and could be used in future studies to look for such trends and correlations.

Wake–sleep and REM–NREM shifts as well as stability measures were computed based on the manually scored hypnogram as well as a data-driven identification of REM, NREM and W stages. A data-driven model recognizes the underlying structure of the data and automatically identifies wake and sleep stages. Using a data-driven sleep staging approach has several advantages over manual scoring.

Please cite this article in press as: Christensen JAE et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.03.006

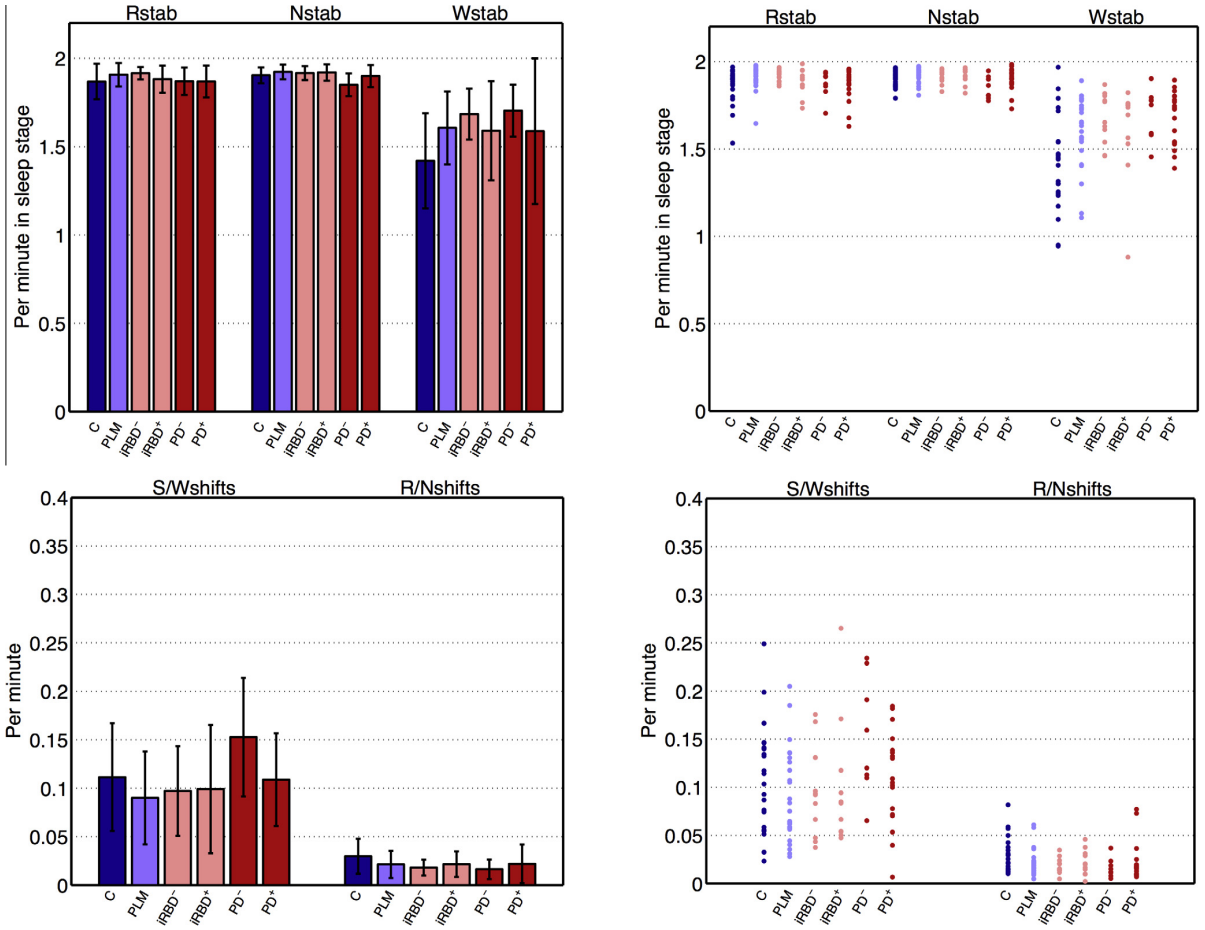


Fig. 2. The frequencies of wake–sleep and REM–NREM shifts and the stability measures for REM sleep, NREM sleep and wake for the six groups based on the manually scored REM, NREM and W epochs. Left, bar charts with heights indicating means, and error bars indicating one standard deviation of the observations above and below the mean. Right, results from each subject indicated as dots. Dark blue, control subjects; light blue, PLMD patients; light red, iRBD patients; dark red, patients with Parkinson's disease (PD). Asterisks indicate a significant between-group difference, determined by Wilcoxon rank-sum tests: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. The Benjamini–Hochberg procedure was used to control the false discovery rate (FDR) at level $q = 0.10$. Fifteen between-group comparisons were made of five measures, giving a total of 75 tests. No comparisons remained significant after FDR correction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

First, automatic staging encompasses micro-sleep characteristics delineated over time intervals as short as one second; the dynamics over such short intervals are not captured by manual scoring, which is of a much lower resolution. The model looks deeper into each sleep epoch and can survey and identify many more sleep characteristics than the human eye can assess. Although final identification is made for 30-s intervals in the automatic detection as well, decisions are based on changes or transitions occurring at a much faster rate. Second, using an automatic, data-driven approach, sleep stages across the night as well as across subjects are more comparable as they are obtained using exactly the same algorithm. Third, using the automatic model, identification of each epoch is based solely on micro-sleep characteristics and changes within the epoch in question, rather than on the scoring of adjacent epochs. As the model features the analysis of short time intervals, it enables a highly detailed description of each sleep epoch, which reduces variation in inter- and intra-subject scoring. Consistent with this, manual scorings for patients with PD have low inter-rater agreement rates (Danker-Hopfe et al., 2004; Jensen et al., 2010).

No significant differences between groups were found for the manually scored hypnograms, whereas the data-driven labeling indicated many between-group differences. The reason for these contradictory results are suggested to be caused by the two different sleep scorings strategies. The fact that the data-driven method makes the labeling solely based on the characteristics of the epoch in question and not the labeling of the prior epoch allows in specific REM sleep to be terminated when no clear EEG and EOG structures supporting REM sleep are present. Contrarily, manually scorings terminate REM sleep only when a clear NREM or W structure is present (Iber et al., 2007). Specifically, REM sleep is besides the characteristic eye movements identified by atonia and EEG similar to that during wakefulness. In periods where eye movements are not present, but the atonia as well as the aroused EEG is maintained, the manual scoring of REM sleep is continued. Relating this to EMG with lack of atonia and EEG with altered characteristics or lack of clear micro-sleep structures as seen in patients with PD or iRBD (Christensen et al., 2014a,b; Dauvilliers et al., 2007; Fantini et al., 2003; Kempfner et al., 2014a,b; Petit et al., 2004), termination of REM sleep is harder to confirm for a manual scorer. As a result, the manually scorings lack to illustrate differences in the REM stability measure, as well as in the REM–NREM sleep transitions.

Please cite this article in press as: Christensen JAE et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.03.006

The sleep regulating mechanisms involve several neurons, which are mainly located in the basal brain regions (Luppi et al., 2011; Saper et al., 2001). Our findings suggest that these mechanisms are affected in patients with iRBD and specifically in patients with PD, suggesting iRBD as an intermediate stage between controls and patients with PD, consistent with Braak's staging theory (Braak et al., 2003). Specifically, REM stability was affected, which involves the neurological networks controlling REM–NREM sleep transitions presented as consisting of REM-on and REM-off areas located in the brainstem (Lu et al., 2006; Luppi et al., 2011). The REM-on area is thought to contain two populations of neurons, where one set projects into the basal forebrain and regulates EEG components of REM sleep, and the other projects into the medulla and spinal cord and regulates atonia during REM sleep (Lu et al., 2006). As a diagnostic criterion of RBD is loss of atonia during REM sleep, the neurons regulating atonia must be affected to some degree in these patients. The REM staging and thereby the REM stability measures in this study were only based on EEG and EOG characteristics, suggesting that the neurons regulating the cortical components of REM sleep are also affected, agreeing with previous findings (Fantini et al., 2003). Specifically, it is proposed that the REM-specific EEG structures are signified to a lesser degree in iRBD and PD patients than in controls and PLMD patients. The reason for this could be neurodegeneration of the ascending branch of the SLD neurons themselves or the heavy innervation between the REM-on and REM-off areas. As the descending branch of the SLD neurons are destroyed in iRBD patients (Luppi et al., 2013, 2011), it is very likely that the ascending branch are affected as well, as these neurons are physically adjacent to each other. In addition, the reported changes in dream enactment in these patients support the hypothesis that the medullo-cortical branch is affected as well.

Another reason for the lower stability of REM sleep in iRBD and PD patients could simply be neurodegeneration affecting prethalamic fibers located in the brain stem. This would cause changes in the patterns, density and/or circadian rhythm of thalamic-induced EEG features such as K-complexes, SS and alpha-activity, etc. Impairment of thalamic control would affect not only the overall sleep rhythm, but also the stability of the different sleep stages.

In a former study we showed that sleep transitions are affected in hypocretin-deficient narcolepsy (Sorensen et al., 2013b). This delineates the role of the hypocretin system in wake–sleep regulation and in NREM–REM regulation. Cerebrospinal hypocretin levels are often normal or subnormal but there are fewer hypocretinergic neurons in hypothalamus in PD (Wienecke et al., 2012). Surprisingly, one study found high a level of hypocretin-1 (orexin-A) (Bridoux et al., 2013), but the small number of patients limited further conclusions. There are limited data from iRBD and PD patients concerning the involvement of hypocretin level. However the spinal hypocretinergic level does not need to fully represent the factual loss of hypocretinergic neurons (Compta et al., 2009), and its involvement in RSWA and the role of the SLD are not fully understood, even in hypocretin-deficient narcolepsy (Knudsen et al., 2010).

It has to be noted that the lower REM sleep stability in PD patients can be due to comorbidities, such as e.g. comorbid insomnia; a common symptom in PD. Riemann et al. (Riemann et al., 2012) reports REM sleep instability as an objective measure for primary insomnia patients compared to good sleeper controls. Although they measure REM sleep instability differently than us (number of manually scored micro-arousals and awakenings per hour of REM sleep compared to our fully data-driven method), we cannot rule out insomnia as causative for the lower REM sleep stability. Insomnia is a diagnosis highly based on subjective measurements, and as we focused on electrophysiological measurements alone, we lack a comparison of the presence and degree of comorbid insomnia to the stability and transitional measurements. Future studies have to address this issue, also including a group of patients with primary insomnia.

A major limitation of this study is that we have not fully investigated the disagreement between the manual and automatic scoring. A weakness of the automatic detector is that it depends on the ability to capture micro-sleep events within the different disease groups. Our "words-in-a-bag" assumption introduces considerable adaptability (each epoch is described by the use of 1192 different "words"), which we believe makes our model capable of better recognizing altered micro-sleep architecture events, such as abnormal SS, K-complexes and eye movements. Optimally, micro-sleep events should be validated and compared in the different groups. However, we do not have a suitable measure of how well the different micro-sleep events are captured. In the case of the control subjects, the manual scores might be a suitable gold standard, although the mean inter-rater agreement rate has been reported to be as low as 76% (Norman et al., 2000). In the patients' case, the inter-rater agreement rates are lower, and in particularly in patients with PD, where the agreement rate has been reported to be only "fair" in more than 25% of sleep recordings (Danker-Hopfe et al., 2004). These results all imply some difficulties with the standard sleep scoring method, whether it be stated as the R&K or the AASM standard. The standards are based on micro- and macro-sleep structures in healthy, young subjects, and it forces the sleep rhythm to be explained by five or six discrete stages defined in 30-s windows. Transitions between sleep stages and from sleep to wakefulness might happen faster than is captured by manual scoring. The fixed period of 30 s is not a physiological parameter, as the brain does not work to a specific or consistent timescale. Analyzing sleep in sleep and/or neurological disorders has proved to be difficult due to disruptions in the sleep transition mechanisms in the brain. Therefore, changes and differences in the micro- and macrostructure of sleep might not be captured by manual scoring with the principles as defined today. The appearance of sleep simply cannot be fitted to the scoring standard.

5. Conclusions

In conclusion, our study suggests that patients with iRBD and PD suffer from instability in the wake–sleep and NREM–REM transitions and instability of wake and sleep. These findings are in accordance with the initial hypothesis that iRBD is an early form of PD and that the basal brain is involved in these diseases in early disease stages. Our results further argue for the problems that these patients suffer from micro-sleep and nocturnal sleep fragmentation. Further studies should be conducted to examine these findings in other neurodegenerative diseases affecting wake and sleep regulation.

Acknowledgments

The PhD project is supported by grants from H. Lundbeck A/S, the Lundbeck Foundation, the Technical University of Denmark and the Center for Healthy Aging, University of Copenhagen. *Conflict of interest:* Julie A. E. Christensen, Henriette Koch, Rune Frandsen, Marielle Zoetmulder, Lars Arvastson, Søren R. Christensen, Helge B. D. Sorensen and Poul Jennum: No conflicts of interest reported.

References

Braak H, Del Tredici K, Rüb U, de Vos RA, Jansen Steur EN, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging 2003;24:197–211.

Bridoux A, Moutereau S, Covali-Noroc A, Margarit L, Palfi S, Nguyen JP, et al. Ventricular orexin-A (hypocretin-1) levels correlate with rapid-eye-movement sleep without atonia in Parkinson's disease. Nat Sci Sleep 2013;12:87–91.

Please cite this article in press as: Christensen JAE et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.03.006

Christensen JAE, Zoetmulder M, Koch H, Frandsen R, Arvastson L, Christensen SR, et al. Data-driven modeling of sleep EEG and EOG reveals characteristics indicative of pre-Parkinson's and Parkinson's disease. J Neurosci Methods 2014a;235:262–76.

Christensen JAE, Kempfner J, Zoetmulder M, Leonthin HL, Arvastson L, Christensen SR, et al. Decreased sleep spindle density in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol 2014b;125:512–9.

Christensen JAE, Koch H, Frandsen R, Kempfner J, Arvastson L, Christensen SR, et al. Classification of iRBD and Parkinson's disease patients based on eye movements during sleep. Conf Proc IEEE Eng Med Biol Soc 2013;2013:441–4.

Compta Y, Santamaria J, Ratti L, Tolosa E, Iranzo A, Muñoz E, et al. Cerebrospinal hypocretin, daytime sleepiness and sleep architecture in Parkinson's disease dementia. Brain 2009;132:3308–17.

Danker-Hopfe H, Kunz D, Gruber G, Klösch G, Lorenzo JL, Himanen SL, et al. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. J Sleep Res 2004;13:63–9.

Dauvilliers Y, Rompré S, Gagnon J, Vendette M, Petit D, Montplaisir J. REM sleep characteristics in narcolepsy and REM sleep behavior disorder. Sleep 2007;30:844–9.

Dos Santos AB, Kohlmeier KA, Barreto GE. Are sleep disturbances preclinical markers of Parkinson's disease? Neurochem Res 2014. http://dx.doi.org/10.1007/s11064-014-1488-7. in press.

Fantini ML, Gagnon JF, Petit D, Rompré S, Décary A, Carrier J, et al. Slowing of electroencephalogram in rapid eye movement sleep behavior disorder. Ann Neurol 2003;53:774–80.

Galvin JE, Lee VM, Trojanowski JQ. Synucleinopathies: clinical and pathological implications. Arch Neurol 2001;58:186–90.

Garcia-Ruiz PJ, Chaudhuri KR, Martinez-Martin P. Non-motor symptoms of Parkinson's disease A review…from the past. J Neurol Sci 2014;338:30–3.

Hansen IH, Marcussen M, Christensen JAE, Jennum P, Sorensen HBD. Detection of a sleep disorder predicting Parkinson's disease. Conf Proc IEEE Eng Med Biol Soc 2013;2013:5793–6.

Iber C, Ancoli-Israel S, Chesson AL, Qua SF. The AASM Manual for the Scoring of Sleep and Associated Events: rules, terminology, and technical specifications. Westchester, IL: American Academy of Sleep Medicine; 2007.

Jennum P, Zoetmulder M, Korbo L, Kjellberg J. The health-related, social, and economic consequences of parkinsonism: a controlled national study. J Neurol 2011;258:1497–506.

Jensen PS, Sorensen HBD, Leonthin HL, Jennum P. Automatic sleep scoring in normals and in individuals with neurodegenerative disorders according to new international sleep scoring criteria. J Clin Neurophysiol 2010;27:296–302.

Kempfner J, Sorensen GL, Nikolic M, Frandsen R, Sorensen HB, Jennum P. Rapid eye movement sleep behavior disorder as an outlier detection problem. J Clin Neurophysiol 2014a;31:86–93.

Kempfner J, Sorensen HB, Nikolic M, Jennum P. Early automatic detection of Parkinson's disease based on sleep recordings. J Clin Neurophysiol 2014b;31:409–15.

Knudsen S, Gammeltoft S, Jennum PJ. Rapid eye movement sleep behaviour disorder in patients with narcolepsy is associated with hypocretin-1 deficiency. Brain 2010;133:568–79.

Koch H, Christensen JAE, Frandsen R, Zoetmulder M, Arvastson L, Christensen SR, et al. Automatic sleep classification using a data-driven topic model reveals latent sleep states. J Neurosci Methods 2014;235:130–7.

Latreille V, Carrier J, Lafortune M, Postuma RB, Bertrand JA, Panisset M, et al. Sleep spindles in Parkinson's disease may predict the development of dementia. Neurobiol Aging 2014. http://dx.doi.org/10.1016/j.neurobiolaging.2014.09.009. in press.

Latreille V, Carrier J, Montplaisir J, Lafortune M, Gagnon JF. Non-rapid eye movement sleep characteristics in idiopathic REM sleep behavior disorder. J Neurol Sci 2011;310:159–62.

Lu J, Sherman D, Devor M, Saper CB. A putative flip-flop switch for control of REM sleep. Nature 2006;441:589–94.

Luppi PH, Clément O, Sapin E, Gervasoni D, Peyron C, Léger L, et al. The neuronal network responsible for paradoxical sleep and its dysfunctions causing narcolepsy and rapid eye movement (REM) behavior disorder. Sleep Med Rev 2011;15:153–63.

Luppi PH, Clément O, Valencia Garcia S, Brischoux F, Fort P. New aspects in the pathophysiology of rapid eye movement sleep behavior disorder: the potential role of glutamate, gamma-aminobutyric acid, and glycine. Sleep Med 2013;14:714–8.

Micallef J, Rey M, Eusebio A, Audebert C, Rouby F, Jouve E, et al. Antiparkinsonian drug-induced sleepiness: a double-blind placebo-controlled study of L-dopa, bromocriptine and pramipexole in healthy subjects. Br J Clin Pharmacol 2009;67:333–40.

Munhoz RP, Teive HA. REM sleep behaviour disorder: How useful is it for the differential diagnosis of parkinsonism? Clin Neurol Neurosurg 2014;127:71–4.

Norman RG, Pal I, Stewart C, Walsleben JA, Rapoport DM. Interobserver agreement among sleep scorers from different centers in a large dataset. Sleep 2000;23:901–8.

Petit D, Gagnon JF, Fantini ML, Ferini-Strambi L, Montplaisir J. Sleep and quantitative EEG in neurodegenerative disorders. J Psychosom Res 2004;56:487–96.

Postuma RB, Gagnon JF, Rompré S, Montplaisir JY. Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease. Neurology 2010;74:239–44.

Riemann D, Spiegelhalder K, Nissen C, Hirscher V, Baglioni C, Feige B. REM sleep instability – a new pathway for insomnia? Pharmacopsychiatry 2012;45:167–76.

Rodrigues Brazète J, Montplaisir J, Petit D, Postuma RB, Bertrand JA, Génier Marchand D, et al. Electroencephalogram slowing in rapid eye movement sleep behavior disorder is associated with mild cognitive impairment. Sleep Med 2013;14:1059–63.

Sakakibara R, Tateno F, Kishi M, Tsuyusaki Y, Terada H, Inaoka T. MIBG myocardial scintigraphy in pre-motor Parkinson's disease: a review. Parkinsonism Relat Disord 2014;20:267–73.

Salawu F, Danburam A, Olokoba AB. Non-motor symptoms of Parkinson's disease: diagnosis and management. Niger J Med 2010;19:126–31.

Saper CB, Chou TC, Scammell TE. The sleep switch: hypothalamic control of sleep and wakefulness. Trends Neurosci 2001;24:726–31.

Saper CB, Fuller PM, Pedersen NP, Lu J, Scammell TE. Sleep State Switching. Neuron 2010;68:1023–42.

Schapira AH, McDermott MP, Barone P, Comella CL, Albrecht S, Hsu HH, et al. Pramipexole in patients with early Parkinson's disease (PROUD): a randomised delayed-start trial. Lancet Neurol 2013;12:747–55.

Schenck CH, Bundlie SR, Mahowald MW. Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behaviour disorder. Neurology 1996;46:388–93.

Schenck CH, Boeve BF, Mahowald MW. Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series. Sleep Med 2013;14:744–8.

Schenck CH, Bundlie SR, Mahowald MW. REM behavior disorder (RBD): delayed emergence of parkinsonism and/or dementia in 65% of older men initially diagnosed with idiopathic RBD, and an analysis of the minimum and maximum tonic and/or phasic electromyographic abnormalities found during REM sleep. Sleep 2003;26:A316.

Schwartz JR, Roth T. Neurophysiology of Sleep and Wakefulness: basic science and clinical implications. Curr Neuropharmacol 2008;6:367–78.

Sixel-Döring F, Trautmann E, Mollenhauer B, Trenkwalder C. Associated factors for REM sleep behavior disorder in Parkinson disease. Neurology 2011;77:1048–54.

Sorensen GL, Mehlsen J, Jennum P. Reduced sympathetic activity in idiopathic rapid-eye-movement sleep behavior disorder and Parkinson's disease. Auton Neurosci 2013a;179:138–41.

Sorensen GL, Knudsen S, Jennum P. Sleep transitions in hypocretin-deficient narcolepsy. Sleep 2013b;36:1173–7.

Sorensen GL, Kempfner J, Zoetmulder M, Sorensen HB, Jennum P. Attenuated heart rate response in REM sleep behavior disorder and Parkinson's disease. Mov Disord 2012;27:888–94.

Stiasny-Kolster K, Mayer G, Schäfer S, Möller JC, Heinzel-Gutenbrunner M, Oertel WH. The REM sleep behavior disorder screening questionnaire-a new diagnostic instrument. Mov Disord 2007;22:2386–93.

Wienecke M, Werth E, Poryazova R, Baumann-Vogel H, Bassetti CL, Weller M, et al. Progressive dopamine and hypocretin deficiencies in Parkinson's disease: is there an impact on sleep and wakefulness? J Sleep Res 2012;21:710–7.

Zoetmulder M, Biernat HB, Nikolic M, Korbo L, Jennum PJ. Sensorimotor gating deficits in multiple system atrophy: comparison with Parkinson's disease and idiopathic REM sleep behavior disorder. Parkinsonism Relat Disord 2014a;20:297–302.

Zoetmulder M, Biernat HB, Nikolic M, Korbo L, Friberg L, Jennum PJ. Prepulse inhibition is associated with attention, processing speed, and 123I-FP-CIT SPECT in Parkinson's disease. J Parkinsons Dis 2014b;4:77–87.

Please cite this article in press as: Christensen JAE et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.03.006

Paper V

TITLE:
Decreased sleep spindle density in patients with idiopathic REM sleep bahavior disorder and patients with Parkinson's disease

AUTHORS:
Julie A. E. Christensen, Jacob Kempfner, Marielle Zoetmulder, Helle L. Leonthin, Lars Arvastson, Søren R. Christensen, Helge B. D. Sorensen and Poul Jennum

JOURNAL:
Clinical Neurophysiology

YEAR:
2014

VOLUME:
125

PAGES:
512-9

PUBLICATION HISTORY:
Submission date: 4 December 2012, acceptance date: 23 August 2013

Clinical Neurophysiology 125 (2014) 512–519

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Decreased sleep spindle density in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease

Julie A.E. Christensen [a,b,*], Jacob Kempfner [a,b], Marielle Zoetmulder [b,c], Helle L. Leonthin [b], Lars Arvastson [e], Søren R. Christensen [e], Helge B.D. Sorensen [a], Poul Jennum [b,d]

[a] Department of Electrical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark
[b] Danish Center for Sleep Medicine University of Copenhagen, Glostrup Hospital, Glostrup, Denmark
[c] Department of Neurology, Bispebjerg Hospital, Copenhagen, Denmark
[d] Center for Healthy Ageing, University of Copenhagen, Copenhagen, Denmark
[e] H. Lundbeck A/S, Copenhagen, Denmark

ARTICLE INFO

Article history:
Accepted 23 August 2013
Available online 11 October 2013

Keywords:
Sleep spindles
Parkinson's disease
REM sleep behavior disorder
Automatic detection
Matching Pursuit
Support Vector Machine

HIGHLIGHTS

- During non-REM sleep, the sleep spindle density is reduced in idiopathic REM sleep behavior disorder (iRBD) and Parkinson's disease (PD) patients, inducing sleep spindles as a potentially PD biomarker.
- This study raises questions on how to validate abnormal sleep spindels in patients with neurodegenerative diseases.
- The introduced new method for automatic sleep spindle detection showed acceptable performance when validated on middle-aged subjects.

ABSTRACT

Objective: To determine whether sleep spindles (SS) are potentially a biomarker for Parkinson's disease (PD).
Methods: Fifteen PD patients with REM sleep behavior disorder (PD + RBD), 15 PD patients without RBD (PD − RBD), 15 idiopathic RBD (iRBD) patients and 15 age-matched controls underwent polysomnography (PSG). SS were scored in an extract of data from control subjects. An automatic SS detector using a Matching Pursuit (MP) algorithm and a Support Vector Machine (SVM) was developed and applied to the PSG recordings. The SS densities in N1, N2, N3, all NREM combined and REM sleep were obtained and evaluated across the groups.
Results: The SS detector achieved a sensitivity of 84.7% and a specificity of 84.5%. At a significance level of $\alpha = 1\%$, the iRBD and PD + RBD patients had a significantly lower SS density than the control group in N2, N3 and all NREM stages combined. At a significance level of $\alpha = 5\%$, PD − RBD had a significantly lower SS density in N2 and all NREM stages combined.
Conclusions: The lower SS density suggests involvement in pre-thalamic fibers involved in SS generation. SS density is a potential early PD biomarker.
Significance: It is likely that an automatic SS detector could be a supportive diagnostic tool in the evaluation of iRBD and PD patients.

1. Introduction

Sleep spindles (SSS) and K-complexes are EEG hallmarks of non-REM (NREM) sleep. SS are generated by a complex interaction between thalamic, limbic and cortical areas and are probably involved in sleep maintenance and memory consolidation (Caporro et al., 2012). These structures are sensitive to involvement in neurodegenerative disorders and it has recently been suggested that changes in SS have the potential to be biomarkers of neurodegenerative disease (NDD) (Ktonas et al., 2009), through the reduced SS activity in patients with Parkinson's disease (PD) (Puca et al., 1973; Myslobodsky et al., 1982; Comella et al., 1993). REM sleep behavior disorder (RBD) is closely associated with PD and is a marker of later

* Corresponding author. Address: Technical University of Denmark, Orsteds plads, building 349, 2800 Kongens Lyngby, Denmark. Tel.: +45 40255689 (mobile), +45 45255737 (work).
E-mail address: julie.a.e.christensen@gmail.com (J.A.E. Christensen).

1388-2457/$36.00
http://dx.doi.org/10.1016/j.clinph.2013.08.013

development of synucleinopathies (Postuma and Montplaisir, 2009; Kempfner et al., 2010). We hypothesize that SS activity is progressively reduced in RBD and PD patients, which would be consistent with the progressive nature of these disorders. SS activity has not previously been evaluated in idiopathic RBD (iRBD). It has been suggested that these patients have a higher risk of developing PD (Schenck et al., 1996), and that this patient group therefore offers an opportunity for detecting early markers of PD before the onset of clinical disease.

Manual scoring of SS is a tedious and time-consuming task. It requires sleep experts and the degree of agreement among experts has been reported as being relatively low: 70 ± 8% (Zygierewicz et al., 1999). Therefore, an automatic SS detector would be valuable for standardizing the SS scores. If SS are potential biomarkers of PD, it could also be a supportive diagnostic tool.

The aims of the study were twofold: firstly, to develop an automatic SS detector based on Matching Pursuit (MP) for feature extraction and a Support Vector Machine (SVM) for classification; secondly, to measure the SS densities found automatically in normal controls, patients with iRBD and PD patients with or without RBD.

2. Methods

2.1. Subjects

Subjects were recruited from patients evaluated at the Danish Center for Sleep Medicine (DCSM) in the Department of Clinical Neurophysiology, Glostrup University Hospital. All patient evaluations included a comprehensive medical and medication history. All patients were assessed by polysomnography (PSG) and with a multiple sleep latency test (MSLT). Patients taking any benzodiazepines, antipsychotic or anti-depressant drug, including hypnotics, were excluded, though dopaminergic treatments were continued. A total of 15 PD patients without RBD (PD − RBD), 15 PD patients with RBD (PD + RBD) and 15 iRBD patients were included. Fifteen age-matched control subjects with no history of movement disorder, dream-enacting behavior or other previously diagnosed sleep disorders were included. Patients using any type of medication known to affect sleep were also excluded.

The demographic data for the three patient groups and the control group are summarized in Table 1.

2.2. Polysomnograph recordings

Polysomnograph (PSG) data were collected in this study. All controls underwent at least one night of PSG recording as outpatients, and all patients underwent at least one night of PSG recording either as outpatients or in hospital in accordance with the AASM standard (Iber et al., 2007). Two or more PSG routines were performed if and only if the prior recording(s) did not meet the technical needs required to make an assessment of acceptable quality. When manually scoring the SS, only the F3-A2, C3-A2 and O1-A2 EEG derivations were visible for the SS scorer, and for 13 control subject a number of randomly selected sleep epochs, each of a duration of 30 s, were chosen for SS scoring. The selection of sleep epochs was carried out by the SS scorer, who aimed at selecting approximately 30 sleep epochs containing one or more visible SS randomly distributed across the sleep cycles. It was ensured that every SS within a chosen sleep epoch was marked. Filter conditions were as stated in the AASM standard, and the AASM standard SS definition was used, whereby SS have frequencies in the range 11–16 Hz, last for 0.5–3 s and have no amplitude criteria. The left EEG derivations were chosen as these are known to exhibit an overall higher spindle density (Bódizs et al., 2009). In order to reproduce realistic conditions, sleep epochs with moderate noise contamination were allowed and no artifacts were removed manually. The scoring yielded a total of 375 sleep epochs with 882 manually scored SS. The distribution of the chosen sleep epochs across the different sleep stages is seen in Table 2. All the scored SS within these sleep epochs were confirmed by an expert.

The raw sleep data, hypnograms and sleep events were extracted from Somnologica Studio (V5.1, Embla, Broomfield, CO 80021, USA) or Nervus (V5.5, Cephalon DK, Nørresundby, Denmark), using the built-in export data tool. For further analysis, the data were imported into MATLAB (R2010b, MathWorks, Inc., Natick, MA, USA).

Table 1
Demographic data for the control and the patient groups.

Patient group	Frequency	Male/female frequency	Age (years)	BMI (kg/m^2)	Sleep efficiency (%)	Bed times (min)
Controls	15	6/9	58.3 ± 9.5	23.2 ± 2.8	88.9 ± 8.4	480 ± 47.5
iRBD	15	12/3	60.1 ± 7.4	24.4 ± 3.1	85.6 ± 8.3	489 ± 95.3
PD − RBD	15	8/7	61.9 ± 6.1	24.7 ± 2.2	82.8 ± 7.9	443 ± 67.2
PD + RBD	15	11/4	62.4 ± 5.2	26.0 ± 3.2	85.4 ± 9.7	445 ± 71.8

2.3. Development of SS detector

The steps in the method for developing the automatic detector are shown in Fig. 1. Firstly, appropriate features were extracted from the C3-A2 and F3-A2 EEG derivations. These are variables that represent characteristics of the classes and may therefore reflect differences between them. These were sent through a classifier that determines the class ('SS' or 'background EEG') to which the data segment belongs.

2.3.1. Feature extraction

Before feature extraction, the polysomnograph C3-A2 and F3-A2 EEG derivations were band pass-filtered from 2 to 35 Hz. The lower cutoff frequency at 2 Hz was chosen to avoid the influence of the high-energy contents at the very low frequencies, and the cutoff at 35 Hz was chosen to reflect the AASM standard.

In this study, the Matching Pursuit (MP) method (Mallat and Zhang, 1993) was chosen for feature extraction in the classification of SS. In the MP signal processing algorithm a given signal is represented by a weighted sum of known basic waveforms, known as Gabor atoms, $g_\gamma(t)$, which in continuous time are expressed as:

$$g_\gamma(t) = K(\gamma)e^{-\pi\left(\frac{t-u}{s}\right)^2}\cos(\omega(t-u)+\phi) \quad (1)$$

Table 2
The distribution of the different sleep stages within the four groups evaluated and for use in the development of the SS detector.

Sleep stage	For use in the development of SS detector	Controls	iRBD	PD − RBD	PD + RBD
Wake (%)	0 (0)	1606 (11)	2220 (15)	2387 (18)	1889 (14)
REM (%)	4 (1)	2710 (19)	2893 (20)	1808 (13)	1761 (13)
N1 (%)	13 (4)	1205 (8)	1238 (8)	1191 (9)	1623 (12)
N2 (%)	330 (88)	6491 (45)	5909 (40)	5817 (44)	5957 (45)
N3 (%)	28 (7)	2388 (17)	2423 (17)	2097 (16)	2128 (16)
Sum (%)	375 (100)	14400 (100)	14683 (100)	13300 (100)	13358 (100)

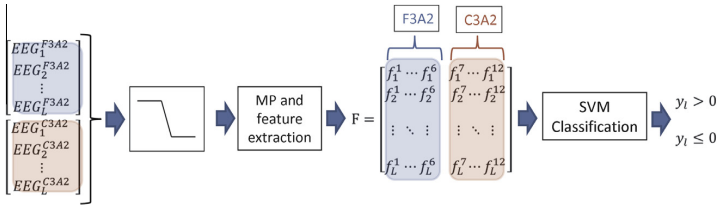


Fig. 1. Method for developing the SS detector. The F3-A2 and C3-A2 EEG derivations are used for feature extraction, divided into L segments of 2 s with 1-s overlap. Before Matching Pursuit and feature extraction, the segments are filtered from 2 to 35 Hz. For each of the L segments, six feature values for each EEG derivation are computed. The feature matrix F of $L \times 12$ features is used as the input for the classification step, which applies a Support Vector Machine and outputs a scalar value y_l for each L segment. The sign of y_l indicates whether the segment corresponds to an SS or not.

Here, $\gamma = \{u, s, \omega, \phi\}$ represents time-shift u and width s in seconds, frequency ω in rad/s and the phase ϕ in rad. $K(\gamma)$ is a normalization scaling factor. By making a redundant dictionary of Gabor atoms, the signal was decomposed iteratively, whereby the Gabor atom most highly correlated with the signal or its residual was chosen at each step. As the iterative process continues, the residual decays exponentially (Mallat and Zhang, 1993), and the process stops when the residual is below a given threshold. The MP algorithm projects a function $f(t)$ on Gabor atoms:

$$f(t) = \sum_{n=0}^{M-1} \langle R^n f(t), g_{\gamma_n}(t) \rangle g_{\gamma_n}(t) + R^M f(t) \tag{2}$$

where g_{γ_0} denotes the first selected atom, $\langle R^n f(t), g_{\gamma_n}(t) \rangle$ is the inner product of the atom and the signal $R^n f(t)$ and $R^M f(t)$ denotes the residual signal after approximating $f(t)$ by using M Gabor atoms.

The time–frequency distribution of the signal energy is derived by adding Wigner–Ville distributions of selected atoms (Mallat and Zhang, 1993), which yields:

$$\begin{aligned} WV_f(t,\omega) = & \sum_{n=0}^{M-1} |\langle R^n f(t), g_{\gamma_n}(t) \rangle|^2 WV_{g_{\gamma_n}}(t,\omega) \\ & + \sum_{n=0}^{M-1} \sum_{k=1, k \neq n}^{M-1} \langle R^n f(t), g_{\gamma_n}(t) \rangle \\ & \times \langle R^k f(t), g_{\gamma_k}(t) \rangle WV_{g_{\gamma_n} g_{\gamma_k}}(t,\omega), \end{aligned} \tag{3}$$

where WV_f and $WV_{g_{\gamma_n}}$ indicate the Wigner–Ville distribution of the signal f and the given Gabor atom g'_{γ_n}, respectively. The first sum corresponds to the auto-terms and the double sum corresponds to the cross-terms of the Wigner–Ville transform. By removing the cross-terms, the energy density of the signal $f(t)$ is found:

$$E_f(t,\omega) = \sum_{n=0}^{M-1} |\langle R^n f(t), g_{\gamma_n}(t) \rangle|^2 WV_{g_{\gamma_n}}(t,\omega). \tag{4}$$

The features were all calculated from the energy densities derived from the Wigner–Ville transform. They were obtained from signal windows of 2 s with a 1-s overlap. For each EEG derivation, the features included:

(1) Three energy features reflecting energy parts in the frequency bands $f < 11$ Hz, $11 \text{ Hz} \leqslant f \leqslant 16$ Hz and $f > 16$ Hz, defining frequencies below, within and above the SS frequency band, respectively.
(2) The logarithm of the energy contribution of the first Gabor atom with a frequency of $11 \text{ Hz} \leqslant f \leqslant 16$ Hz.
(3) The logarithm of the maximum energy point in the energy density found by Eq. (4) and the corresponding frequency.

The six feature values were calculated for the C3-A2 and F3-A2 EEG derivations, yielding a total of 12 feature values for each 2-s segment. The features were normalized with respect to the 95th percentile of the features, since this was the normalization method found to perform best.

Further information about the MP algorithm and details about the implementation of the original software written in C into MATLAB, which is used in this study, can be found in Mallat and Zhang (1993) and Mallat et al. (2008).

2.3.2. Classification

In this study, the Support Vector Machine (SVM) algorithm (Boser et al., 1992; Cortes and Vapnik, 1995) was chosen to classify the SS. SVM is a binary supervised learning method, and has proved to be efficient when dealing with datasets of unequal size. Clearly, the essential goal in all machine learning techniques is to optimize the generalized classification properties of the model, i.e. to categorize correctly as many data points of an unseen dataset as possible. This optimization process is employed in the training phase, and the essence of SVM is to find optimal separating hyperplanes in a high-dimensional feature space (Cristianini and Shawe-Taylor, 2000). The optimization in SVM consists of maximizing the margin between classes in the feature space, which is sometimes referred to as "the maximal margin classifier" (Cristianini and Shawe-Taylor, 2000).

A training dataset can mathematically be described as:

$$\{x_i, y_i\}_{i=1}^{L} \quad y_i \in \{-1, 1\} \quad x_i \in \Re^D, \tag{5}$$

where each of the L training samples x_i is a vector with D feature values and y_i takes the value of -1 or 1, indicating the group to which each training sample i belongs. In the case of the two classes being linearly separable, they can be classified by a hyperplane described as:

$$h(x_i) = \langle x_i, w \rangle + b = 0, \tag{6}$$

where w is the normal to the hyperplane and b is a shifting constant. The finding of the hyperplane is based on the positive and negative samples of x (y_l in Fig. 1) that are most strongly indicative of the slope of the resulting separating hyperplane. These are the support vectors, and they all satisfy the constraint:

$$y_i \cdot (\langle x_i, w \rangle + b) - 1 + \xi_i \geqslant 0 \quad \forall i, \tag{7}$$

where $\xi_i \geqslant 0$ $\forall$i is a slack variable introducing a cost or penalty to misclassified samples, relaxing the constraints of the fully linearly separable case. The penalty increases with the distance to the separating hyperplane.

To describe the separating hyperplane, the values for w and b are found by solving the problem summarized to:

$$\begin{cases} \min(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{L} \xi_i) \\ y_i(\langle x_i, w \rangle + b) - 1 + \xi_i \geqslant 0 \quad \forall i \\ \xi_i \geqslant 0 \quad \forall i \end{cases} \tag{8}$$

where the cost parameter C is a user-defined parameter indicating the penalty for misclassification. The problem is solved by introduc-

ing Lagrange multipliers, and knowing the values for w and b defines the optimal orientation of the separating hyperplane, and the SVM classifier is defined. The classification of a new unknown data point $x' = [f^1 \ \ldots \ f^{12}]$ indicated by the 12 features described above merely requires the sign of the function:

$$h(x') = \langle x', w \rangle + b \tag{9}$$

to be evaluated. The sign indicates on which side of the separating hyperplane the data point x' lies.

The SVM classification can easily be extended to work on non-linear separable classes by using kernels $K(x_i,x_j)$, mapping the data into a Euclidean space H where they can be linearly separated. In this study, a Radial Basis Function (RBF) kernel was used for the SVM, and a parameter optimization study was performed by doing a grid search on the cost parameter C and the kernel-specific parameter $\gamma = \frac{1}{2\sigma^2}$, which controls the flexibility of the decision boundaries with higher γ values allowing greater flexibility. The evaluated values were $\gamma = \{0.125, 0.25, 0.5, 1, 2, 3\}$ and $C = \{1, 4, 16, 64, 256, 1024\}$. The optimal pair for the final model was found to be $(C, \gamma) = (256, 1)$.

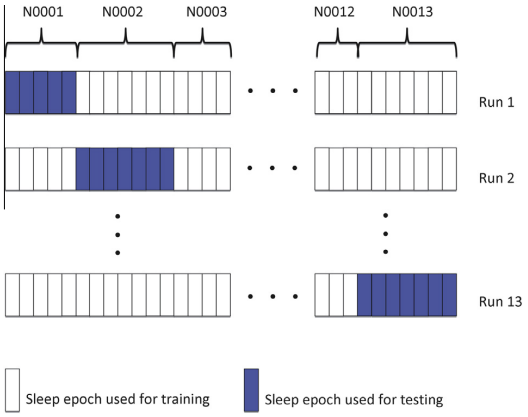


Fig. 2. Illustration of the leave-one-subject-out strategy used in this study. Each small rectangle represents a sleep epoch. Blue and white rectangles are used for testing and training, respectively. The numbers N0001–N0013 are the IDs for the control subjects. Different numbers of sleep epochs were available from each subject, so different amounts of data were held out in each run.

As in other studies, only the data with manually scored SS was used in the development of the automatic SS detector (Schönwald et al., 2006; Causa et al., 2010). Hence, the feature vectors from the sleep epochs with manual scores of SS were used to train and test the classifier in this study. Each second of EEG data was labeled either SS (1) or background EEG (−1). The training and testing phases employed the leave-one-subject-out strategy. As illustrated in Fig. 2, the test data set in each of the 13 runs were of unequal size, as the number of available scored sleep epochs differed between the control subjects. Overall performance measures were calculated as the mean of the 13 runs.

The SVM^{perf} algorithm developed by Thorsten Joachims at Cornell University was used in this study (Joachims, 2005, 2006; Joachims and Yu, 2009). This can be found at (Joachims, 2009).

3. Results

3.1. Performance of automatic SS detector

To validate the performance of the algorithm, different statistical measures were defined on the basis of four variables: True Positives (TP), False Positives (FP), True Negatives (TN) and False

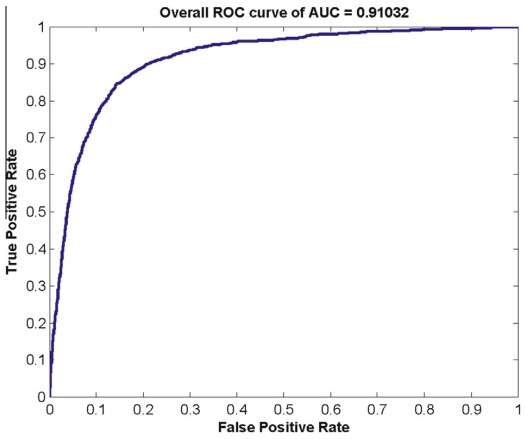


Fig. 4. The overall ROC curve for a mean AUC measure of 91.0%, based on the leave-one-subject-out method.

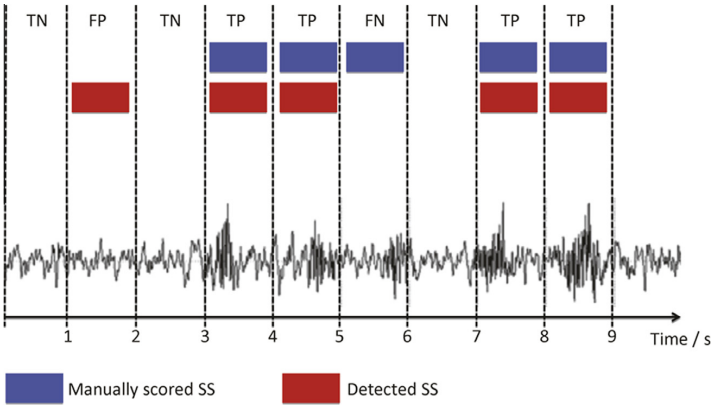


Fig. 3. Definition of the four variables, True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN), based on seconds.

Negatives (FN). These were found by comparing the SS detected by the algorithm and those manually scored, as illustrated in Fig. 3.

The values obtained were used to calculate the sensitivity and specificity, and by using these, a Receiver Operating Characteristics (ROC) curve was derived (Fawcett, 2006) (Fig. 4). These values were obtained using the data with manually scored SS, i.e. the epochs stated under "For use in the development of SS detector" in Table 2.

The area under the ROC curve (AUC) reached 91.0% based on the leave-one-subject-out strategy. By choosing the (FP, TP) pair as the point on the ROC curve, where the sign of the function described in Eq. (10) determined the class, the mean sensitivity reached 84.7% and the mean specificity reached 84.5%. These were considered satisfactory for the purpose of this study.

3.2. SS densities

To determine whether the SS density varied between the three groups of patients and the control group, the automatic detector was applied to the all-night recordings from lights-off until lights-on. The total number and the distribution of the different sleep stages within the four groups are provided in Table 2. SS density was defined as SS/min and measured for the different sleep stages. Specifically, sleep epochs of N1, N2, N3, all NREM and REM were evaluated separately. The values of the means and standard deviations of the various sleep stages and groups are shown in Table 3.

To establish whether there was a significant difference between the means of SS density in the four groups, unpaired two-sample *t*-tests were performed. The variances within each group were assumed to be unequal. Comparisons of the control group with a diseased group used one-sided *t*-tests in order to establish whether the mean of the control group was higher than those of each of the diseased groups. Comparisons of pairs of diseased groups used two-sided *t*-tests in order to establish whether the means of the diseased groups differed from one another. The significant differences are illustrated in Fig. 5.

At a significance level of $\alpha = 1\%$, the iRBD and PD patients with RBD had a significantly lower mean SS density than the control group in N2, N3 and all NREM combined. At a significance level of $\alpha = 5\%$, the PD patients without RBD had a significantly lower mean SS density than the control group in N2 and all NREM combined. No significant differences were found for the results obtained in REM sleep for either group.

4. Discussion

We describe a new method for automatic SS detection combining MP for feature extraction and SVM for classification with acceptable performance. There were three main findings in this study. First, the group of PD patients with RBD had a significantly lower SS density than the control group in N2, N3 and all NREM combined. Also, iRBD patients had a significantly lower SS density than controls in N2, N3 and all NREM combined. These results were significant at the level of $\alpha = 1\%$, but by relaxing this to a value of $\alpha = 5\%$, the PD patients without RBD also showed a significantly lower SS density compared with controls in N2 and all NREM combined.

Table 3
Means and standard deviations of the SS densities of the four groups in the respective sleep stages. SS density was defined as SS/min.

Sleep stage	N1	N2	N3	All NREM	REM
Controls	4.4 ± 1.6	6.2 ± 1.5	5.6 ± 1.3	6.0 ± 1.3	2.2 ± 1.4
iRBD	4.4 ± 1.7	4.7 ± 1.9	4.1 ± 2.4	4.5 ± 1.8	2.8 ± 1.4
PD − RBD	4.4 ± 1.7	5.1 ± 1.8	4.9 ± 2.3	5.0 ± 1.5	2.4 ± 1.4
PD + RBD	4.4 ± 2.1	4.2 ± 1.9	3.6 ± 2.1	4.2 ± 1.8	3.6 ± 2.2

The automatically obtained SS densities for the four groups lie slightly above the interval expected on the basis of previous findings (Emser et al., 1988; Happe et al., 2004; Durka et al., 2005b). (Emser et al., 1988) reported only SS densities obtained by manual scorings in N2, and for the control subjects they found values of 5.8 ± 1.2 SS/min, compared with 6.2 ± 1.5 SS/min obtained in the present study. They included 12 control subjects with an average age of 53.4 ± 8.6 years, compared with the 15 control subjects with an average age of 58.3 ± 9.5 years included in this study. In a more recent study, however, the manually found SS densities were reported to be 100.2 ± 52.3 SS in one hour of N2 sleep obtained from 10 controls with an average age of 65.2 ± 10.7 years (Happe et al., 2004). Again, although the mean ages slightly differ, the foundations of the studies are comparable, and in this case, the results were much lower than those in the present study. We therefore conclude that it is indeed difficult to define a normal range for SS density, and while the results obtained in the present study are within the range found by some studies, they are outside the range reported in others.

The significant differences in SS densities between the PD patients and the control subjects are consistent with the findings of (Emser et al., 1988), but contradict those of (Happe et al., 2004). However, both these studies only included findings from N2 sleep and did not separate the PD patients into those with and without RBD. To the best of our knowledge, no other studies of SS density in PD patients divide the patients in this way. Furthermore, no studies detecting SS densities in iRBD patients are known. These key points make this study unique, because all sleep stages and the influence of RBD are investigated. The patient groups involved in our study provide the best basis for investigating whether the SS density is a possible biomarker for PD.

The results obtained in this study suggest that the more diseased a subject is, the lower is their SS density. Considering PD patients with RBD to belong to the most diseased group, it is hypothesized that the iRBD group and the PD group without RBD are intermediate stages between the control group and the most diseased group. More specifically, our results do not contradict the general theory that iRBD is an intermediate stage between healthy subjects and PD patients.

There are two hypotheses that could explain the results of the present study. The first originated from a study performed by (Braak et al., 2003), who studied how PD was related to the intraneuronal pathology progress and, on this basis, proposed a staging procedure of incidental and symptomatic PD cases. They argued that the disease process in the brain stem follows an ascending course with little interindividual variation, starting with the involvement of basal brain structures (Braak stage 1 and 2) and spreading to the other brain structures (Braak stage 3 and 4) (Braak et al., 2003). This indicates that the progressive course of the neurodegeneration in PD at some point reaches the SS generator (in the thalamus) located above the "start point" of the disease.

Patients with iRBD had a lower SS density than control subjects in both N2, N3 and all NREM sleep combined. For all cases, the iRBD group had a lower p-value than the PD group without RBD. The neurodegeneration probably did not yet involve the SS generator in the thalamus, but influenced the prethalamic afferent fibers (PRE) arising from the brainstem and posterior hypothalamus. These fibers interacted with the thalamocortical pyramidal cells (TC), which are thought to generate SS by interacting with the thalamic reticular neurons (RE). This suggests that the lower SS densities seen in stages intermediate to PD might be due to affected or defective PRE fibers, indicating a PD in very early stages (Braak stage 1 and 2). This also suggests that the SS density might be more

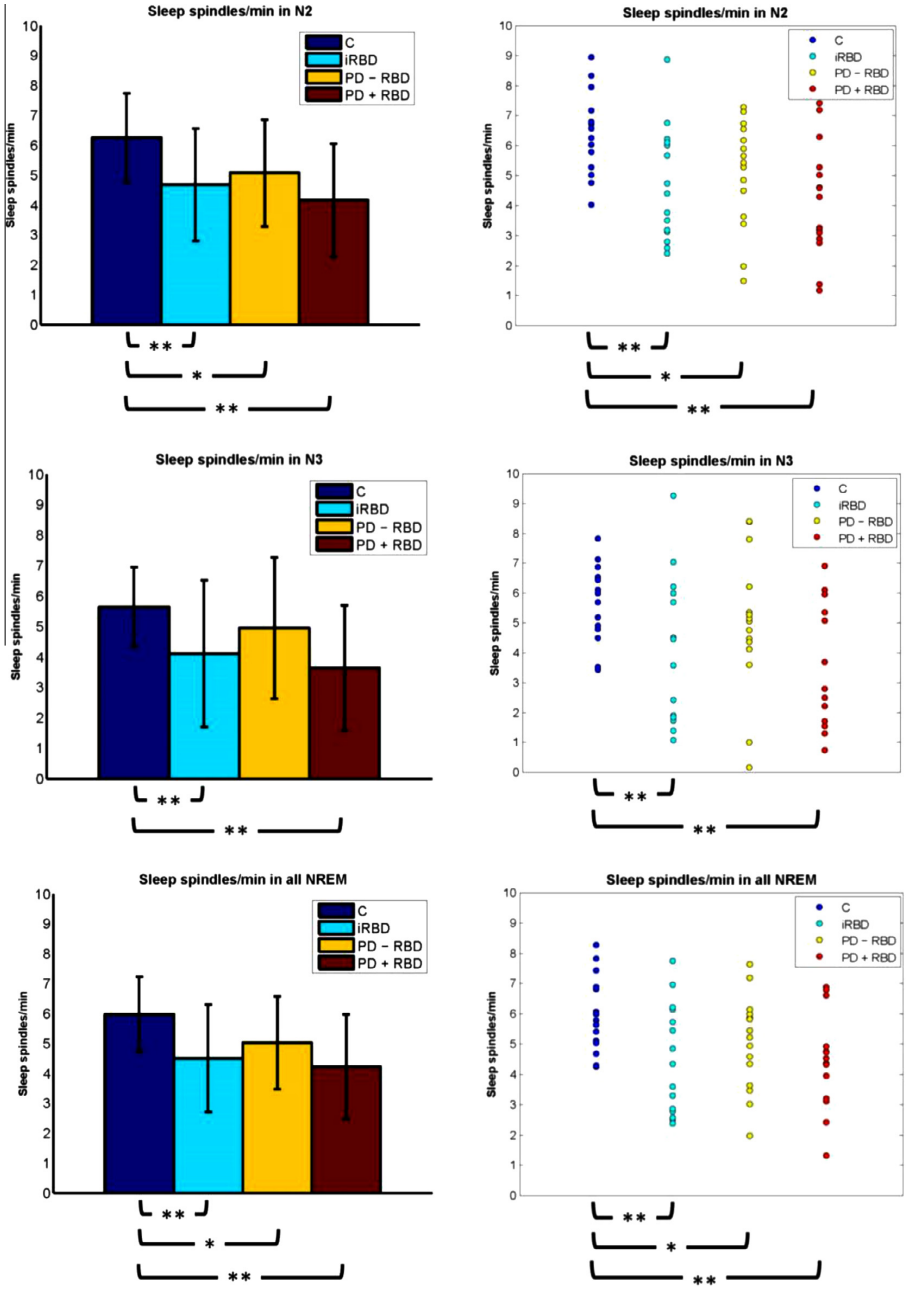


Fig. 5. Results for N2, N3 and all NREM combined. The figures on the left illustrate the mean and standard deviation of the individuals in the four groups, while the individual measures for each subject and patient are seen in the figures on the right. A single asterisk indicates significant changes with $p < 0.05$. Double asterisks indicate significant changes with $p < 0.01$.

affected by RBD as a disease than by PD as a disease. Following this hypothesis, this could explain why other studies investigating SS densities in neurodegenerative diseases do not find significant changes (Happe et al., 2004; Rauchs et al., 2008). In the study by Rauchs et al., 2008, significant differences in spindle types, but no differences in mean SS intensity could be found when comparing patients with Alzheimer's Disease (AD) to an age-matched unmedicated control group.

Several studies have reported links between consolidation of memory and increases of SS density during sleep that followed learning (Steriade and Timofeev, 2003; Schmid et al., 2012). Also, the influence of different pharmacological interventions on SS den-

sity has been investigated (Puca et al., 1973; Myslobodsky et al., 1982). Puca et al., 1973 reported no differences in SS density between L-dopa treated PD patients and control subjects, suggesting that dopamine has a positive effect on the amount of SS. All PD patients in this study were treated in accordance to clinical standards as optimally as possible, and dopaminergic treatment was not discontinued in treated patients. Consequently a dose response was not possible to obtain in these patients. However, we doubt whether this would show a causal effect as the spindle activity is dependent on prethalamic innervation and thalamic function. Effect of dopaminergic drugs depends on the clinical DOPA response, which decreases with disease advancement and potentially also with reduced spindle activity. This will call for further studies including patients with different disease severities. Additionally, the study of (Happe et al., 2004) concluded that the number of SS did not correlate with the dopamine transporter binding in the striatum or the serotonin binding in the thalamic/hypothalamic region. Supplementary, the present study finds significant decreased amount of SS in the group of untreated iRBD patient compared to control subjects. The various dopaminergic treatments of the PD patients enrolled in the present study were continued, and no information of the activities prior to sleeping is included in this study. However, the sleep efficiency and the total recording time (TRT) is noted for the four groups in Table 1, and although the PSG parameters differ, there was no association between the SS densities and the PSG parameters. Finally, it should be pointed out that besides the reported changes in different SS parameters caused by age, pre-PSG parameters, disease, medications etc., the great inter-individual variability makes it difficult to investigate or state the effect of the different influences.

The automatic SS detector developed in this study performs similarly to other detectors described in the literature (Devuyst et al., 2006; Schönwald et al., 2006; Causa et al., 2010). These studies used other definitions of SS. We used the AASM standard, and to our knowledge, no other studies have evaluated SS detection using this definition. Furthermore, previous studies used younger subjects who tended to have SS with a more prominent amplitude and density (Guazzeli et al., 1986; Nicolas et al., 2001; Happe et al., 2004). We used older subjects and applied the method on PSG obtained from iRBD an PD patients. In developing the detector, we used a very limited extract of data and the training was carried out primarily with N2 and to a lesser degree with N1, N3 and REM sleep. As seen in Table 2, the distribution of sleep stages with manual SS scorings is not comparative with the distribution of sleep stages from the complete nights of sleep for the four groups. However, the SS densities were drawn from complete nights of sleep, including all epochs from lights-off to lights-on, and it is therefore a possible concern that not enough data was available to train with this algorithm, in general and within the different sleep epochs.

To refine the algorithm further, more manually SS scorings are clearly needed, and additional diseased PSG should be included. This could also provide an insight into how well the detector performs for the diseased groups, which could be worse than for the control subjects. A major issue is how to validate abnormal SS in patients with neurodegenerative diseases. Two questions arise: (1) should these be compared with normal, age-standardized SS, or (2) should validation be against a separate SS score in EEG with SS of abnormal configuration?

A key matter emerging from this study is whether it is reasonable to conclude on the basis of results obtained by an automatic detector with a sensitivity of 84.5% and a specificity of 84.7%. In the worst case, the significant differences in the present study could have arisen from this underlying uncertainty in the SS results.

Several studies have considered the accuracy of manually scored SS (Zygierewicz et al., 1999; Durka et al., 2005a,b). Also, some studies have shown that their performance measures improve when raising an amplitude threshold (Durka et al., 2005a; Causa et al., 2010). Additionally, the rate of agreement of two independent sleep scorers has been found to be as low as 70 ± 8% (Zygierewicz et al., 1999). Considering these points together does indeed call into question the validity of human judgement when scoring SS. A feature extraction algorithm such as MP makes it possible to identify SS buried in noise or very closed spaced in time or frequency. Other algorithms based on MP or other feature extraction algorithms have shown that many of the erroneous detections are mostly confined to the epochs containing artifacts (Durka et al., 2005b; Causa et al., 2010). It can be debated whether training an algorithm to perform "as well" as a human sleep scorer really does produce a better SS detector or forces it to miss the buried or atypical SS. When relating this problem to EEGs from diseased subjects, these atypical SS could be of great interest as they might contain information that would never be spotted by the human eye.

5. Conclusion

The study develops a novel approach for designing an automatic SS detector that was considered a satisfactory method for standardizing the SS scorings. Applying this detector to data from iRBD and PD patients as well as age-matched controls, SS densities were obtained from different sleep stages and proved to be significantly lower for the iRBD group and the PD groups with and without RBD compared with the controls in NREM sleep. The lower SS density suggests involvement in pre-thalamic fibers involved in SS generation. We conclude that SS is a potential biomarker for PD, and it is likely that an automatic SS detector could be a supportive diagnostic tool. Further research is needed into both the automatic detection of SS and the evaluation of the SS changes in early RBD and PD detection.

Acknowledgements

The Ph.D. project is supported by Grants from H. Lundbeck A/S, the Technical University of Denmark and Center for Healthy Ageing, University of Copenhagen, Copenhagen, Denmark. Julie A.E. Christensen, Jacob Kempfner, Helle L. Leonthin, Lars Arvastson, Søren R. Christensen, Helge B.D. Sorensen and Poul Jennum: No conflicts of interest reported.

References

Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory; 1992. p. 144–52.

Braak H, Del Tredici K, Rüb U, de Vos RA, Jansen SEN, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging 2003;24:197–211.

Bódizs R, Körmendi J, Rigó P, Lázár AS. The individual adjustment method of sleep spindle analysis: methodological improvements and roots in the fingerprint paradigm. J Neurosci Methods 2009;178:205–13.

Caporro M, Haneef Z, Yeh HJ, Lenartowicz A, Buttinelli C, Parvizi J, et al. Functional MRI of sleep spindles and K-complexes. Clin Neurophysiol 2012;123:303–9.

Causa L, Held CM, Causa J, Estévez PA, Perez CA, Chamorro R, et al. Automated sleep-spindle detection in healthy children polysomnograms. IEEE Trans Biomed Eng 2010;57:2135–46.

Comella CL, Tanner CM, Ristanovic RK. Polysomnographic sleep measures in Parkinson's disease patients with treatment-induced hallucinations. Ann Neurol 1993;34:710–4.

Cortes C, Vapnik V. Support-vector networks. In: Cortes C, Vapnik V, editors. Machine learning 1995;vol. 20. Netherlands: Springer; 1995. p. 273–97. no. 3.

Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other Kernel-based learning methods. Cambridge University Press; 2000.

Devuyst S, Dutoit T, Didier JF, Meers F, Stanus E, Stenuit P, et al. Automatic sleep spindle detection in patients with sleep disorders. Conf Proc IEEE Eng Med Biol Soc 2006;1:3883–6.

Durka PJ, Malinowska U, Szelenberger W, Wakarow A, Blinowska KJ. High resolution parametric description of slow wave sleep. J Neurosci Methods 2005a;147:15–21.

Durka PJ, Matysiak A, Montes EM, Sosa PV, Blinowska KJ. Multichannel matching pursuit and EEG inverse solutions. J Neurosci Methods 2005b;148:49–59.
Emser W, Brenner M, Stober T, Schimrigk K. Changes in nocturnal sleep in Huntington's and Parkinson's disease. J Neurol 1988;235:177–9.
Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006;27:861–74.
Guazzeli M, Feinberg I, Aminoff M, Fein G, Floyd TC, Maggini C. Sleep spindles in normal elderly: comparison with young adult patterns and relation to nocturnal awakening, cognitive function and brain atrophy. Electroencephalogr Clin Neurophysiol 1986;63:526–39.
Happe S, Anderer P, Pirker W, Klösch G, Gruber G, Saletu B, et al. Sleep microstructure and neurodegeneration as measured by [123I]beta-CIT SPECT in treated patients with Parkinson's disease. J Neurol 2004;251:1465–71.
Iber C, Ancoli-Israel S, Chesson AL, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events: rules, terminology, and technical specifications. Am. Acad. Sleep Med. 2007.
Joachims T. A support vector method for multivariate performance measures. Proc Int Conf Mach Learn 2005;119:377–84.
Joachims T. Training linear SVMs in linear time. In: Proceedings of the ACM SIGKDD international conference on knowledge discovery an data mining; 2006. p. 217–26.
Joachims, T. Support vector machine for multivariate performance measures [Guide to SVMperf algorithm]. September 7, 2009. Available at: <http://www.cs.cornell.edu/People/tj/svm_light/svm_perf.html> [accessed 16.05.12].
Joachims T, Yu C-NJ. Sparse kernel SVMs via cutting-plane training. Mach Learn 2009;76:179–93.
Kempfner J, Sorensen G, Zoetmulder M, Jennum P, Sorensen HB. REM behaviour disorder detection associated with neurodegenerative diseases. Conf Proc IEEE Eng Med Biol Soc 2010:5093–6.
Ktonas PY, Golemati S, Xanthopoulos P, Sakkalis V, Ortigueira MD, Tsekou H, et al. Time-frequency analysis methods to quantify the time-varying microstructure of sleep EEG spindles: possibility for dementia biomarkers? J Neurosci Methods 2009;185:133–42.
Mallat SG, Zhang Z. Matching pursuits with time–frequency dictionaries. IEEE Trans Signal Process 1993;41:3397–415.
Mallat SG, Zhang Z, Franaszczuk PJ, Jouny CC. Matching Pursuit Software. August 21, 2008. Available at: <http://erl.neuro.jhmi.edu/mpsoft/> [accessed 16.05.12].
Myslobodsky M, Mintz M, Ben-Mayor V, Radwan H. Unilateral dopamine deficit and lateral EEG asymmetry: sleep abnormalities in hemi-Parkinson's patients. Electroencephalogr Clin Neurophysiol 1982;54:227–31.
Nicolas A, Petit D, Rompré S, Montplaisir J. Sleep spindle characteristics in healthy subjects of different age groups. Clin Neurophysiol 2001;112:521–7.
Postuma RB, Montplaisir J. Predicting Parkinson's disease – why, when, and how? Parkinsonism Relat Disord 2009;15:105–9.
Puca FM, Bricolo A, Turella G. Effect of L-DOPA or amantadine therapy on sleep spindles in Parkinsonism. Electroencephalogr Clin Neurophysiol 1973;35:327–30.
Rauchs G, Schabus M, Parapatics S, Bertran F, Clochon P, Hot P, et al. Is there a link between sleep changes and memory in Alzheimer's disease? Neuroreport 2008;19:1159–62.
Schenck CH, Bundlie SR, Mahowald MW. Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder. Neurology 1996;46:388–93.
Schmid MR, Murbach M, Lustenberger C, Maire M, Kuster N, Achermann P, et al. Sleep EEG alterations: effects of pulsed magnetic fields versus pulse-modulated radio frequency electromagnetic fields. J Sleep Res 2012;21:620–9.
Schönwald SV, de Santa-Helena EL, Rossatto R, Chaves ML, Gerhardt GJ. Benchmarking matching pursuit to find sleep spindles. J Neurosci Methods 2006;156:314–21.
Steriade M, Timofeev I. Neuronal plasticity in thalamocortical networks during sleep and waking oscillations. Neuron 2003;37:563–76.
Zygierewicz J, Blinowska KJ, Durka PJ, Szelenberger W, Niemcewicz S, Androsiuk W. High resolution study of sleep spindles. Clin Neurophysiol 1999;110:2136–47.

Paper VI

TITLE:
Sleep spindle alterations in patients with Parkinson's disease

AUTHORS:
Julie A. E. Christensen, Miki Nikolic, Simon C Warby, Henriette Koch, Marielle Zoetmulder, Rune Frandsen, Keivan K Moghadam, Helge B. D. Sorensen, Emmanuel Mignot and Poul Jennum

JOURNAL:
Frontiers in Human Neuroscience

YEAR:
2015

VOLUME:
9

PAGES:
Article no 233: 1-13

PUBLICATION HISTORY:
Submission date: 16 December 2014, acceptance date: 11 April 2015

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 01 May 2015
doi: 10.3389/fnhum.2015.00233

Sleep spindle alterations in patients with Parkinson's disease

Julie A. E. Christensen[1,2,3]*, *Miki Nikolic*[2], *Simon C. Warby*[4], *Henriette Koch*[1,2,3], *Marielle Zoetmulder*[2,5], *Rune Frandsen*[2], *Keivan K. Moghadam*[6], *Helge B. D. Sorensen*[1], *Emmanuel Mignot*[3] and *Poul J. Jennum*[2,7]

[1] *Biomedical Engineering, Department of Electrical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark,* [2] *Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Glostrup University Hospital, Glostrup, Denmark,* [3] *Stanford Center for Sleep Sciences and Medicine, Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA, USA,* [4] *Center for Advanced Research in Sleep Medicine, Sacré-Coeur Hospital of Montréal, University of Montréal, Montréal, QC, Canada,* [5] *Department of Neurology, Bispebjerg Hospital, Copenhagen, Denmark,* [6] *Department of Biomedical and Neuromotor Sciences (DIBINEM), University of Bologna, Bologna, Italy,* [7] *Center for Healthy Ageing, University of Copenhagen, Copenhagen, Denmark*

The aim of this study was to identify changes of sleep spindles (SS) in the EEG of patients with Parkinson's disease (PD). Five sleep experts manually identified SS at a central scalp location (C3-A2) in 15 PD and 15 age- and sex-matched control subjects. Each SS was given a confidence score, and by using a group consensus rule, 901 SS were identified and characterized by their (1) duration, (2) oscillation frequency, (3) maximum peak-to-peak amplitude, (4) percent-to-peak amplitude, and (5) density. Between-group comparisons were made for all SS characteristics computed, and significant changes for PD patients vs. control subjects were found for duration, oscillation frequency, maximum peak-to-peak amplitude and density. Specifically, SS density was lower, duration was longer, oscillation frequency slower and maximum peak-to-peak amplitude higher in patients vs. controls. We also computed inter-expert reliability in SS scoring and found a significantly lower reliability in scoring definite SS in patients when compared to controls. How neurodegeneration in PD could influence SS characteristics is discussed. We also note that the SS morphological changes observed here may affect automatic detection of SS in patients with PD or other neurodegenerative disorders (NDDs).

Keywords: Parkinson's disease, sleep spindle morphology, EEG, neurodegeneration, biomarker

OPEN ACCESS

Edited by:
Christian O'Reilly, McGill University, Canada

Reviewed by:
Ki-Young Jung, Korea University Medical Center, South Korea
Géraldine Rauchs, Institut National de la Santé et de la Recherche Médicale, France
Veronique Latreille, Hopital du Sacré-Coeur de Montréal, Canada

***Correspondence:**
Julie A. E. Christensen, Technical University of Denmark, Orsteds Plads, Building 349, DK-2800 Kongens Lyngby, Denmark
julie.a.e.christensen@gmail.com

Received: *16 December 2014*
Accepted: *11 April 2015*
Published: *01 May 2015*

Citation:
Christensen JAE, Nikolic M, Warby SC, Koch H, Zoetmulder M, Frandsen R, Moghadam KK, Sorensen HBD, Mignot E and Jennum PJ (2015) Sleep spindle alterations in patients with Parkinson's disease. Front. Hum. Neurosci. 9:233. doi: 10.3389/fnhum.2015.00233

Introduction

Parkinson's disease (PD) is a neurodegenerative disorder (NDD) characterized primarily by motor symptoms, including bradykinesia, rigidity, postural instability, and tremor. Although the disease process in PD is not restricted to a specific brain area, these symptoms are mostly caused by the loss of dopaminergic neurons in the substantia nigra pars compacta resulting in a reduction or depletion of dopamine (Galvin et al., 2001). Lewy body aggregations of alpha-synuclein in the brain are a central feature of PD pathology (Galvin et al., 2001). These inclusions typically start in caudal areas of the brain and progress anteriorly (Braak et al., 2003), and may take place years prior to involvement of the substantia nigra and associated development of motor symptoms.

Abbreviations: AASM, American Academy of Sleep Medicine; EEG, electroencephalography; iRBD, idiopathic REM sleep behavior disorder; MSA, Multiple System Atrophy; NDD, Neurodegenerative disorders; PD, Parkinson's disease; PSG, polysomnographic; REM, Rapid eye movements; SS, Sleep spindles.

Specifically, Braak et al.'s PD staging is based on Lewy-body distribution, which rise from the dorsal motor nucleus of the vague nerve in the medulla and in the olfactory bulb (stage 1) emerging through the subceruleus-ceruleus complex and the magnocellularis reticular nucleus (stage 2), the substantia nigra, the pedenculopontine nucleus and the amygdala (stage 3), the temporal mesocortex (stage 4), and finally reaching the neocortex (stage 5 and 6). Stage 1 and 2 were considered as pre-Parkinsonian states, stage 3 and 4 as Parkinsonian states and 5 and 6 as late-Parkinsonian states (Braak et al., 2003).

In addition to the motor manifestations that define PD, non-motor symptoms such as sleep problems, depression, dementia and attention deficit (Chaudhuri et al., 2011, 2006), autonomic symptoms as abnormal heart rate variability (Sorensen et al., 2012, 2013) and gastrointestinal symptoms such as nausea and constipation (Garcia-Ruiz et al., 2014) are all well known in patients with PD. Stating the presence of at least two of the four motor symptoms resting tremor, bradykinesia, rigidity, and postural imbalance typically makes the clinical diagnosis of PD, although it has been indicated that the pathological changes in the striatal dopaminergic system develop several years before the clinical appearance of PD. Further development of the pathology may result in Lewy Body Dementia.

Twenty years ago, it was discovered that idiopathic rapid eye movement (REM) sleep behavior disorder (iRBD) is closely related to Parkinsonism (Schenck et al., 1996, 2013a; Salawu et al., 2010). Indeed, the presence of iRBD, even without the presence of motor or cognitive complaints, confers a significant risk of conversion into synnucleinopathies including PD (Iranzo, 2011; Schenck et al., 2013b). The diagnosis of RBD requires complaints or an anamnesis describing dream enactment behaviors as well as a manifestation of REM sleep without atonia (RSWA) as measured by polysomnography (PSG) (Stevens and Comella, 2013; American Academy of Sleep Medicine, 2014). The idiopathic form of RBD (iRBD) is diagnosed when no concurrent neurological disease is found, and International classification of Sleep Disorders criteria for RBD are met (Stevens and Comella, 2013; American Academy of Sleep Medicine, 2014). Specifically, measures of RSWA (Postuma et al., 2010; Kempfner et al., 2013), slow wave characteristics (Latreille et al., 2011), sleep stability and differences in electroencephalographic (EEG) or electrooculographic micro- and macro-sleep patterns have been investigated in patients with iRBD and/or PD (Christensen et al., 2012, 2013, 2014b).

Reduced sleep spindle (SS) density and activity have been identified in patients with PD and iRBD (Puca et al., 1973; Myslobodsky et al., 1982; Emser et al., 1988; Comella et al., 1993; Christensen et al., 2014a; Latreille et al., 2015). SS are generated by a complex interaction involving thalamic, limbic, and cortical areas. A di-synaptic circuit between thalamic reticular neurons and thalamocortical relay cells, both located in the thalamus, can spontaneously generate spindle-like oscillations, which are conveyed to the cortex by the axons of the thalamocortical relay cells. These cells receive feedback from cortical pyramidal cells as well as input from pre-thalamic fibers originating from the brainstem and posterior hypothalamus (Steriade et al., 1993; Steriade and Timofeev, 2003). As such the thalamus holds a primary role in generating and controlling SS. SS have been reported to have a gating role with regard to the flow of thalamic sensory input, and thus may have a sleep-preserving role (De Gennaro and Ferrara, 2003). Also, several studies have reported SS to have an important role in memory consolidation, synaptic plasticity and cognition (Steriade and Timofeev, 2003; Schabus et al., 2006; Fogel and Smith, 2011; Fogel et al., 2012; Latreille et al., 2015). The formation of SS begins in the infant brain (De Gennaro and Ferrara, 2003), but SS characteristics such as density and amplitude change with age (Nicolas et al., 2001; De Gennaro and Ferrara, 2003), suggesting that SS play an important role in normal cognitive functioning.

Although a reduction in SS density is not specific to PD, SS and other EEG features may be potential useful as biomarkers of disease progression or therapeutic efficacy in PD and other NDDs (Nguyen et al., 2010; Leiser et al., 2011; Micanovic and Pal, 2014). However, the identification of SS is a difficult task; studies assessing inter-scorer variance in normal sleep have shown significant variance in SS identification, both between human experts and between automated SS detectors (Warby et al., 2014; Wendt et al., 2014). SS identification and characterization in pathological sleep is not well studied, but previous evidence suggests that SS may have different characteristics in PD patients (Latreille et al., 2015), and therefore may interfere with traditional sleep staging in patients (Comella et al., 1993; Jensen et al., 2010; Christensen et al., 2014b; Koch et al., 2014).

In this study, we aimed to identify changes in SS density and specific morphological characteristics of SS in patients with PD. Since five sleep experts identified SS independently, we were also able to assess inter-expert variation of SS identification in EEG of patients and controls. By identifying specific changes in SS characteristics, we aimed to better understand the mechanism and to what extent the neurodegenerative progress influences SS characteristics, also identifying specific spindle features that may be useful as prognostic biomarkers of disease. A secondary aim was to help guide the specialized development of automatic SS detectors to be used on EEG from patients with NDDs.

Materials and Methods

Subjects and Recordings

Polysomnographic (PSG) EEG data from 15 patients with PD and 15 sex- and age-matched control subjects with no history of movement disorder, dream-enacting behavior or other previously diagnosed sleep disorders were included in this study. The subjects were all recruited from the Danish Center for Sleep Medicine (DCSM) in the Department of Clinical Neurophysiology, Glostrup University Hospital in Denmark. All patients were evaluated by a movement specialist with a comprehensive medical and medication history and a PSG analyzed according to the American Academy of Sleep Medicine (AASM) standard (Iber et al., 2007). The diagnostic certainty for PD at Danish neurological departments has been reported to be 82% (Wermuth et al., 2012). None of the PD patients had dementia at inclusion, but one of the patients with PD later developed Multiple System Atrophy (MSA), indicated as the Parkinsonian type (MSA-P) as the patient had predominating

PD-like symptoms. Subjects were excluded from the study if they were taking medications known to effect sleep (antidepressants, antipsychotics, hypnotics). However, dopaminergic treatments were permitted despite their potential effect on vigilance and SS characteristics (Puca et al., 1973; Micallef et al., 2009). In addition to ethical concerns regarding discontinuing dopaminergic treatment in these subjects, we wanted to avoid deleterious discontinuation effects on the PSG, as well as unpleasant and negative motor effects that could interfere with the study. The quality of each PSG recording was individually examined, and recordings with disconnections or significant amounts of signal artifact were not included. Demographic data and PSG variables for the two groups are seen in **Table 1**.

Manual Labeling of Sleep Spindles

For each subject, eight blocks of five consecutive epochs of non-REM sleep stage 2 (N2) of 30-s duration were selected randomly from the PSG recording in between lights off and lights on. The blocks were randomly chosen and ranked by use of Matlab's *randsample*-function. One-by-one and in the prioritized order, the blocks were visually checked for major movements or other contaminating artifacts. The first eight artifact-free blocks were chosen as the ones to be scored for SS. A total of five independent sleep experts identified SS in these blocks, where only the C3-A2 EEG derivation was visible. The signals were filtered with a notch filter at 50 Hz and a band-pass filter with cutoff frequencies at 0.3 Hz and 35 Hz, as indicated by AASM standards (Iber et al., 2007). All analyzed signals had a sampling frequency of 256 Hz. The experts assigned a confidence score to each identified spindle, to indicate the amount of confidence in the identification (as described previously in Warby et al., 2014). In this way, each SS was given a confidence weighting of 1 for "definitely SS," 0.75 for "probably a SS" and 0.5 for "maybe a SS."

The scoring procedure was performed in a Matlab-based software program "EEG viewer" developed by MN at DCSM. The program mimics a standard sleep scoring program in a clinical setting, and includes the standard features so the experts have the same opportunities to view and navigate the PSG data as they are used to when analyzing sleep in the clinic. The program ensures that if an epoch to be scored does not have any marked SS, the expert is required to click a box saying "no spindles in current epoch." This ensures that the total of 40 epochs of N2 sleep per subject was analyzed by each expert. The experts were blinded for which group the subjects belong to.

TABLE 1 | Demographic and PSG data for the two groups studied.

Characteristics	PD patients	Controls	*P*
Total counts (Male/Female)	15 (7/8)	15 (7/8)	–
Age (Years)	62.7 ± 5.8	62.9 ± 5.9	0.90
BMI (kg/m^2)	25.3 ± 3.5	22.1 ± 2.5	0.02
Disease duration (years)	6.7 ± 4.5	NA	–
Hoehn and Yahr stage	2.0 ± 1.2	NA	–
UPDRS part III "on"	20.9 ± 7.0	NA	–
ACE	90.2 ± 4.8	NA	–
Levodopa equivalent dosage (mg)	621.1 ± 301.5	NA	–
Levodopa use [n (%)]	10 (67)	NA	–
Dopamine agonist use [n (%)]	14 (93)	NA	–
Sleep efficiency (%)	79.7 ± 14.1	87.1 ± 8.4	0.09
Time in bed (min)	448.1 ± 82.0	499.6 ± 63.7	0.07
LM index (number/hour)	31.8 ± 34.8	30.4 ± 35.3	0.91

BMI, Body Mass Index; UPDRS, Unified Parkinson's disease rating scale; ACE, Addenbrooke's cognitive examination; LM, Leg movements.

The final SS identifications used for morphology measures were defined using the group consensus rule described in Warby et al. (2014). Spindle identifications from five different experts with weighted confidence scores for each SS were averaged at each sample point and aggregated into a single consensus. Sample points that had an average score of higher than the group consensus threshold $T_{gc} = 0.25$ were included in the final group consensus, and the morphology measures were computed on these group consensus SS. It was decided to use $T_{gc} = 0.25$ as this was found to be the best in Warby et al. (2014).

Spindle Characteristics and between Group Comparisons

The morphology of the identified SS was characterized by their (1) duration, (2) oscillation frequency, (3) maximum peak-to-peak amplitude, (4) percent-to-peak amplitude, and (5) SS density per minute; all of which are well-evaluated elsewhere (Warby et al., 2014). The morphology measures were all computed using Matlab 2013b. Before any of the measures were computed, the central EEG signal was filtered forward and reverse with (1) a notching filter with the notch at 50 Hz and a bandwidth of 50/35 Hz (at −3 dB) and (2) a 4th order Butterworth band-pass filter with cut off frequencies (−3 dB) at 0.3 Hz and 35 Hz.

For each SS the duration was computed in seconds as

$$dur = \frac{\#\ samples}{f_s},$$

where $f_s = 256$ Hz is the sampling frequency and *# samples* defines the number of samples. The samples were consecutive and obeyed the consensus rule. The oscillation frequency was defined in Hz and was for each SS estimated as

$$f_{osc} = \frac{K}{2 \cdot dur},$$

where K defines the total number of extrema points detected using Matlab's *findpeaks*-function applied on a 5-point moving average smoothed version of the SS signal and with a minimum peak-to-peak distance of 11 samples. The maximum points were found by applying the *findpeaks*-function directly, and the minima points were found by applying the function on the flipped signal, and the total number of extrema points was set as the sum of the two. These settings were chosen, as they were considered best for estimating the f_{osc} when visually investigating numerous randomly selected examples of SS. The maximum peak-to-peak amplitude was for each SS estimated as

$$A_{p2p} = \max\left(\left|A_e\left(k+1\right) - A_e(k)\right|\right), k = 1, 2, \ldots, K-1,$$

where A_e is a vector holding the amplitude values for each of the K detected extrema points. To investigate the influence on SS from K-complexes or delta waves, the maximum peak-to-peak amplitude was estimated twice for each SS; once without any further frequency filtering of the data, and once where the data was forward and reverse filtered with a 10th order high-pass filter with cut off frequency (−3 dB) at 4 Hz to remove low frequency, high amplitude waves that may interfere with the peak-to-peak calculation. The percent-to-peak amplitude gives a simple measure between 0 and 1 of the symmetry of the spindle and it was computed for each SS as

$$Sym = \frac{\#\ samples\ before\ point\ of\ A_{p2p}}{\#\ samples},$$

where the point of A_{p2p} is defined as the point between the maxima and minima delineating A_{p2p}. Finally, the density was computed for each subject as the number of SS per minute of investigated data, described as

$$Density = \frac{2 \cdot \#\ SS}{\#\ epochs\ reviewed}.$$

The morphology measures were computed for the SS identifications for each expert, as well as for the spindles included in the group consensus. For the SS included in the group consensus, a minimum duration threshold $dur_{th} = 0.2$ s was used, and resulted in the exclusion of only three spindles. This threshold is less that the minimum duration stated by the AASM scoring (0.5 s). However, others have shown that apparent spindles <0.5 s are clearly recognizable by sleep experts, and have similar characteristics to spindles >0.5 s (Warby et al., 2014). We used a minimum duration threshold of 0.2 s because we wanted to determine whether PD patients and controls have specific differences in these shorter spindles. When computing the measures for the SS identifications for each expert, all the SS were included, regardless of their confidence score and duration. Two-sided Wilcoxon rank sum tests with a significance level of $\alpha = 0.05$ were used for each of the measures to test for significant differences between the two groups.

Inter-Expert Reliability When Scoring SS

Inter-expert reliability measures were computed for each of the 10 available expert-pairs. True positives (TP) define the number of samples where both experts have marked SS, true negatives (TN) define the number of samples where both experts have not marked SS, false positives (FP) define the number of samples where the reference-expert has not marked SS, and the other expert has and false negatives (FN) define the number of samples where the reference-expert has marked SS, but the other expert has not. For each comparison, the reliability measures were indicated as the F_1-score and the Cohen's Kappa coefficient (κ). The F_1-score is the harmonic mean of precision (P) and recall (R) and reaches its best value at 1 (perfect agreement) and the worst at 0 (no agreement). It is computed as

$$\begin{aligned} F_1\text{-score} &= \frac{2 \cdot R \cdot P}{R + P}, \text{ where} \\ R &= \frac{TP}{TP + FN} \text{ and } P = \frac{TP}{TP + FP}. \end{aligned}$$

The κ is often used to measure inter-annotator reliability as it takes the agreement occurring by chance into account. It reached its best value at 1 (perfect agreement) and worst at -1 (no agreement). It reaches 0 when accuracy is equal to what is expected by chance. It is computed as

$$\begin{aligned} \kappa &= \frac{\frac{TP+TN}{N} - Pr}{1 - Pr}, \text{ where} \\ Pr &= \frac{TP + FN}{N} \cdot \frac{TP + FP}{N} + \\ &\quad \left(1 - \frac{TP + FN}{N}\right) \cdot \left(1 - \frac{TP + FP}{N}\right), \end{aligned}$$

where $N = TP + TN + FP + FN$ defines the total number of samples reviewed. The relative strength of agreement associated with κ can been described by the labels "poor" ($\kappa < 0.00$), "slight" ($0.00 \leq \kappa \leq 0.20$), "fair" ($0.21 \leq \kappa \leq 0.40$), "moderate" ($0.41 \leq \kappa \leq 0.60$), "substantial" ($0.61 \leq \kappa \leq 0.80$) and "almost perfect" ($0.81 \leq \kappa \leq 1.00$) (Landis and Koch, 1977). The F_1-score and κ are symmetric regarding false detections and will therefore both yield the same regardless of which expert were used as the reference.

Results

For the SS included in the group consensus, it was found that patients with PD show SS that are significantly different from controls in terms of duration, oscillation frequency and max peak-to-peak amplitude. Additionally, patients with PD have significantly different SS density compared to controls. Specifically, it was found that patients with PD have decreased SS density (−38.17%/−0.71 SS/min), and that their SS are longer (+11.69%/+0.09 s), have a lower frequency (−2.27%/−0.29 Hz) and higher max peak-to-peak amplitude (+19.61%/9.45 μV) compared to controls (**Table 2**). No significant differences were identified for the symmetry measure. The maximum peak-to-peak amplitude estimated after removal of frequencies below 4 Hz was still significantly different between groups. Of note, patients with PD still showed a higher max peak-to-peak amplitude (+20.95%/9.49 μV) compared to controls. The five SS morphology measures are illustrated in **Figure 1**. From left to right, the eight first ID numbers in both groups are females ranging from the youngest to the oldest. The last seven IDs in both groups are males, also ranging from the youngest to the oldest. One of the patients later developed MSA and is illustrated with black.

The patients had significantly fewer spindles than the controls (p-value < 0.05). Ten patients and only four controls had less than 10 SS in the 40 epochs of N2 sleep that were assessed; four

TABLE 2 | Mean (μ) and standard deviation (σ) for the spindle characteristics found for each of the experts' identifications as well as for the spindles in the group consensus.

Spindle characteristic	Expert 1 (947 SS)		Expert 2 (752 SS)		Expert 3 (952 SS)		Expert 4 (282 SS)		Expert 5 (2135 SS)		Group consensus (901 SS)		*P*
	PD	*C*	*PD*	*C*	*PD*	*C*	*PD*	*C*	*PD*	*C*	*PD*	*C*	
Duration [sec, μ ± σ]	0.93 ± 0.44	0.84 ± 0.41	0.66 ± 0.29	0.67 ± 0.29	0.74 ± 0.29	0.68 ± 0.27	0.88 ± 0.20	0.77 ± 0.24	1.19 ± 0.52	1.12 ± 0.51	0.86 ± 0.35	0.77 ± 0.36	$<0.002^{a,c,d,e,GC}$
Frequency [Hz, μ ± σ]	12.38 ± 1.27	12.69 ± 1.27	12.92 ± 1.24	13.07 ± 1.11	12.45 ± 1.22	12.62 ± 1.34	12.73 ± 1.14	13.13 ± 1.00	11.69 ± 1.24	12.03 ± 1.28	12.51 ± 1.21	12.80 ± 1.23	$<0.02^{a,b,d,e,GC}$
Max peak-to-peak amplitude [μV, μ ± σ]	57.37 ± 17.23	48.26 ± 15.37	56.96 ± 18.14	46.88 ± 15.96	57.75 ± 17.15	48.52 ± 15.47	64.60 ± 16.68	49.95 ± 14.04	51.44 ± 18.14	45.02 ± 15.73	57.64 ± 17.34	48.19 ± 15.55	$<0.001^{a,b,c,d,e,GC}$
Max peak-to-peak amplitude, after removal of frequencies < 4 Hz [μV, μ ± σ]	53.87 ± 15.99	45.24 ± 14.17	54.38 ± 16.85	44.16 ± 14.89	54.89 ± 16.34	45.43 ± 14.18	62.40 ± 16.64	47.51 ± 13.13	46.20 ± 16.62	40.15 ± 13.92	54.78 ± 16.24	45.29 ± 14.41	$<0.001^{a,b,c,d,f,GC}$
Percent-to-peak amplitude [μ ± σ]	0.49 ± 0.23	0.47 ± 0.24	0.46 ± 0.23	0.46 ± 0.23	0.46 ± 0.24	0.46 ± 0.23	0.46 ± 0.22	0.46 ± 0.21	0.45 ± 0.25	0.45 ± 0.25	0.47 ± 0.23	0.46 ± 0.23	NS
Density [per min, μ ± σ]	1.29 ± 2.44	1.87 ± 1.56	0.91 ± 1.36	1.60 ± 1.27	1.16 ± 1.95	2.01 ± 1.82	0.30 ± 0.51	0.64 ± 0.84	2.91 ± 2.52	4.21 ± 2.14	1.15 ± 2.06	1.86 ± 1.57	$<0.05^{a,b,c,GC}$

Wilcoxon rank sum tests were used to test for significance between patients with Parkinson's disease (PD) and control subjects (C). [a]significant for expert 1, [b]significant for expert 2, [c]significant for expert 3, [d]significant for expert 4, [e]significant for expert 5, [GC]significant for group consensus.

patients and 0 controls had no SS. Only 3 patients compared to 10 controls had more than 20 SS in the group consensus.

As a supplementary check, the significance tests were performed on SS identifications from each of the five experts individually. The maximum peak-to-peak amplitude was, for all five experts, both before and after removal of frequencies below 4 Hz, significantly different in patients with PD compared to controls. The duration and oscillation frequency were also significantly different between the two groups for 4/5 of the experts, and density significantly different between the two groups for 3/5 of the experts. The mean and standard deviations of the SS morphology measures and the results from the significance tests are summarized in **Table 2**.

Figure 2 illustrates the relation between the SS measures and disease duration for the patients, and **Figure 3** illustrates the relation between the SS measures and Addenbrooke's Cognitive Examination (ACE) score for the patients. Note that the x-axes are not continuous, but denote disease duration in years (**Figure 2**) and ACE score (**Figure 3**) for 15/15 and 13/15 of the patients, respectively. The three subjects with highest SS density are all females, and the one with the highest SS density is a patient with PD later diagnosed with MSA-P (indicated as PD+MSA in the figures). She is illustrated with black in **Figures 1**, **2**, **3**. No clear visual tendency between SS characteristics and disease duration or ACE score was seen for any of the measures. **Supplementary Figure 1** illustrates the relation between SS measures and Hoehn and Yahr (H and Y) stage and **Supplementary Figure 2** illustrates the relation between SS measures and the Unified Parkinson's Disease Rating Scale (UPDRS) Part III. No clear visual trends were seen.

Considering that the outlier PD patient with a very high spindle density (highest of all subjects in the study) later developed MSA, we reanalyzed the SS included in the group consensus when results from this outlier patient were left out, and found the same measures to be as significant different between the groups. Specifically, patients now have an even bigger decrease in SS density (−61.29%/−1.14 SS/min), a longer SS duration (+11.69%/+0.09 s), a slower frequency (−4.14%/−0.53 Hz) and a higher max peak-to-peak amplitude, both before (+16.93%/8.16 μV) and after (+17.95%/8.13 μV) removal of low frequencies when compared to controls. The results for this analysis are summarized in **Table 3**.

Figure 4 shows scatterplots for the individual SS, where the maximum peak-to-peak amplitude (before removal of low frequencies) defines the y-axis and the oscillation frequency and duration defines the x-axis, respectfully. Linear trend lines are added on top of the scatterplots in order to see differences between groups. We found a trend of a positive correlation between the duration and maximum peak-to-peak amplitude. Interestingly, SS from patients showed this tendency to a lesser degree (slope of +11.74 μV/s) compared to SS from controls (slope of +18.09 μV/s). Also, we found a negative correlation of oscillation frequency and maximum peak-to-peak amplitude, and found this tendency to be less apparent for SS from patients (slope of −1.02 μV/Hz) compared to SS from controls (slope of −4.10 μV/Hz).

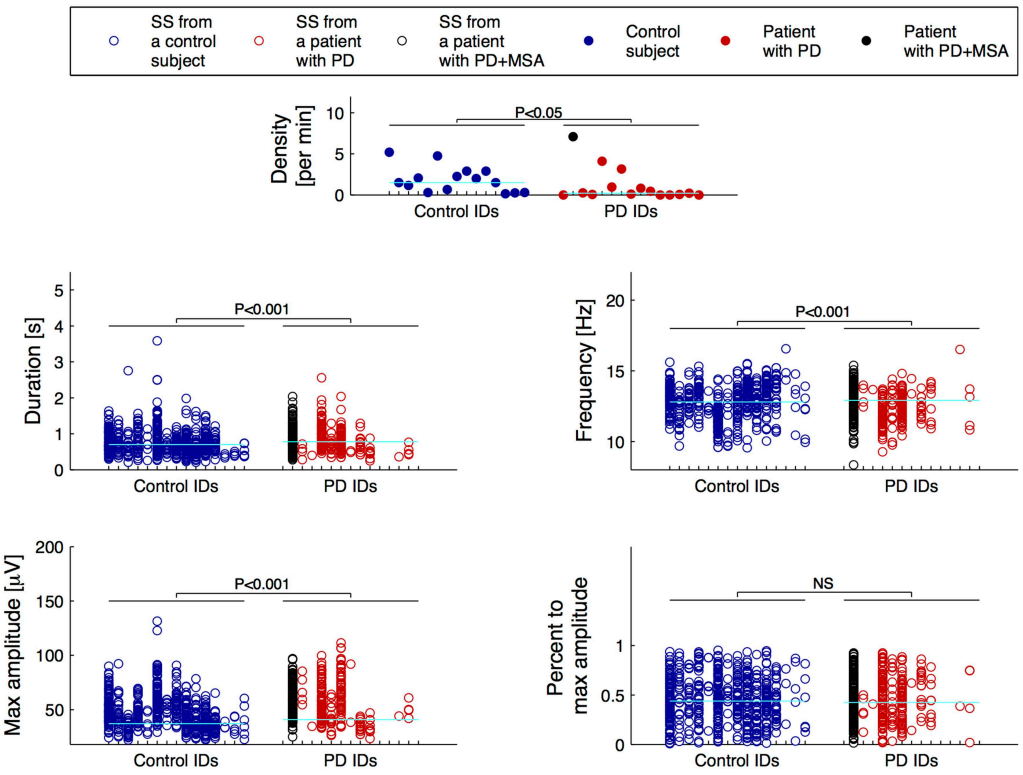


FIGURE 1 | Distributions of the morphology measures for the spindles included in the group consensus. From left to right, the first eight IDs in both groups are females ranging from the youngest to the oldest, and the following seven IDs are males also ranging from the youngest to the oldest. One patient with Parkinson's disease (PD) later developed Multiple System Atrophy (MSA) and is indicated with black. The cyan horizontal lines indicate the group median for each of the measures.

Table 4 summarizes the fraction of SS included in the group consensus that do not strictly pass AASM criteria for a spindle (11–16 Hz, 0.5–3.0 s). Overall, 25.3% of the SS identified by experts and included in the group consensus did not meet AASM criteria. Most of these "abnormal" SS would have been excluded because their duration is too short (16.9%) or have an oscillation frequency that is too slow (9.7%).

In order to determine if there was a difference between PD and controls in the frequency of "abnormal" spindles not meeting AASM criteria, we compared the groups. All 15/15 control subjects had SS, whereas only 11/15 patients with PD had some SS. It was found that control subjects show significantly more "abnormal" spindles not meeting AASM criteria, i.e., more spindles with a too short duration compared to patients with PD (**Table 4**). No significant difference was however found between groups when the outlier patient with PD + MSA was left out of the analysis.

When computing the SS characteristic based on AASM criteria, the same SS characteristics were found to be significantly different between PD patients and controls (**Table 5**). Analysis of these SS showed that patients with PD have a decreased density (−32.84%/−0.44 SS/min), and their SS are longer (+9.41%/+0.08 s), have a lower frequency (−2.69%/−0.35 Hz) and higher max peak-to-peak amplitude before removal of low frequencies (+21.34%/+10.37 μV) and after (+22.51%/+10.30) compared to controls. These differences are similar to those found based on all SS in the group consensus.

Table 6 summarizes inter-expert reliabilities of SS scoring, where the SS are grouped according to their confidence score. The mean inter-expert reliability of scoring "definite SS" computed by κ was found to be significant lower for patients compared to controls. Although not significant, a trend for a lower κ was found for "probable/definite SS" in patients compared to controls ($P = 0.054$). In all cases, the inter-expert reliability is lower for scoring SS in patients compared to controls.

Discussion

Based on a group consensus of manually scored SS from five independent sleep experts, this study investigates morphological changes of SS in a central EEG lead of patients with PD compared to age- and sex-matched control subjects. The main findings of this study are that patients with PD have a decreased SS density, and that their SS have a longer duration, a slower oscillation

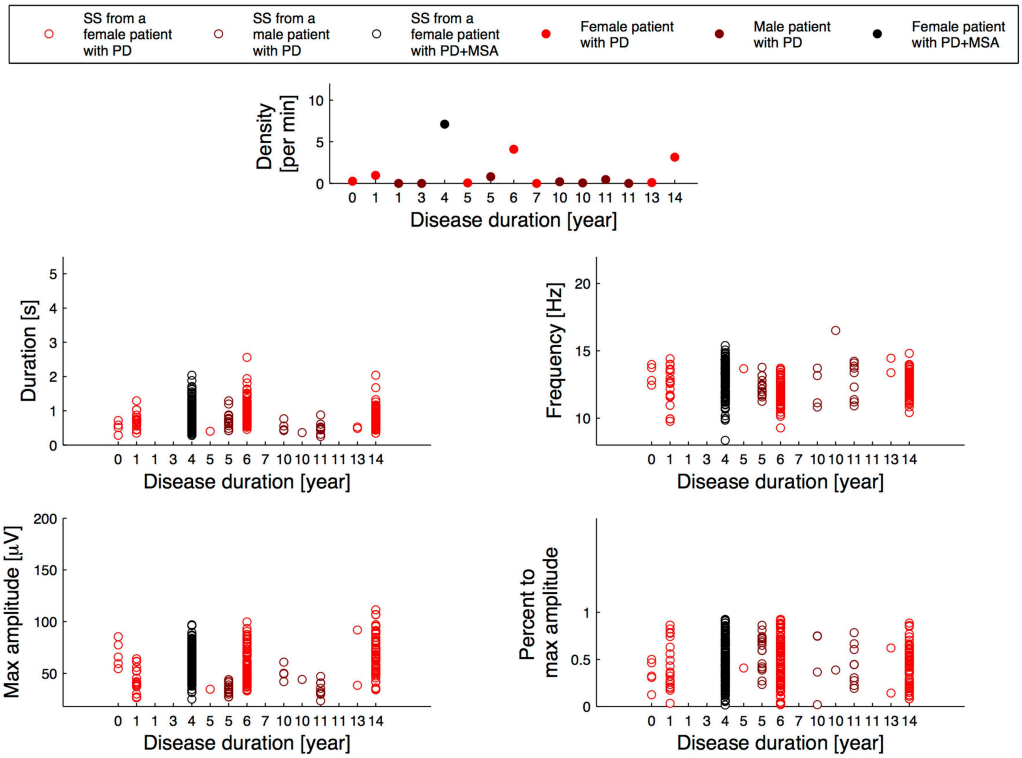


FIGURE 2 | Distribution of the morphology measures for the spindles from patients with Parkinson's disease (PD), where the patients are sorted according to their disease duration. PD+MSA indicates a patient with PD, that later developed Multiple System Atrophy (MSA).

frequency and higher maximum peak-to-peak amplitude. These results suggest that not only SS density but also specific morphological changes in SS have potential clinical utility when diagnosing PD. Further, the data suggests that the disease process affect directly or indirectly the brain regions responsible for the generation of SS. Future studies including more subtypes of PD and NDDs in general are however needed to investigate whether the specific morphological changes in SS can be used to differentiate different PD subtypes as well as different NDDs.

The results illustrate the fact that there are fewer SS in patients with PD, and that the few that are remaining are more pronounced when compared to those seen in controls. There could be several explanations for this. First, patients with PD have a more "blurred" EEG in general with either a lack of or an abnormal mixture of micro- and macro-sleep structures (Petit et al., 2004; Christensen et al., 2014b). This pattern may make it more difficult to identify distinct SS, as they would be buried within other undefined EEG microstructural changes. In this case, only the obvious SS would rise over background and be marked. Second, it could be that the neurodegenerative process has affected the thalamic neurons responsible for generating and controlling SS in such a way, that SS are only generated when very strong signals from pre-thalamic fibers reaches the thalamus resulting in more pronounced SS. Third, we cannot rule out that these SS changes could be the result of treatment with dopaminergic agents affecting the morphology of SS, although a previous report suggests that these drugs should increase spindle density (Puca et al., 1973), which is not what we observed.

It was found that patients with PD have a lower SS density compared to age and sex-matched controls. This finding is consistent with our and other groups' prior findings (Emser et al., 1988; Christensen et al., 2014a; Latreille et al., 2015), but contradicts those of other studies (Happe et al., 2004). According to Braak et al. (2003), the neurodegenarative progress in PD shows a progressive ascending course starting from the brain stem and spreading to additional brain structures. At some point, the neurodegeneration may affect or destroy the SS generator of the thalamus, resulting in fewer or no spindles. Interestingly, (Roth et al., 2000) found that medial thalamotomy abolishes spindle activity in N2 sleep systematically, but that pallido-thalamic tractotomy attenuate spindle activity only to a varying degree, with spindles reemerging after 3 months. It is therefore likely that neurodegenerative involvement of prethalamic fibers from the brain stem may affect spindle activity to a certain degree. In **Figure 1**, it is apparent that for four of the patients, no SS are included in the group consensus, and that for six other patients, less than 10 spindles were identified.

Surprisingly, a PD patient showing an abnormally high SS density was later diagnosed with MSA-P. Although only a single case, it is an interesting finding which support the hypothesis that spindles can be used as a marker of diagnostic subgroups of PD. Latreille et al. (2015) reported a decline in SS activity paralleling cognitive decline in patients with PD, suggesting that SS activity could be used as an early marker of Dementia. The number of patients included in present study is, however, too small to perform further subgroup analysis. Additionally, in both groups, younger subjects and females trend in showing slightly higher spindle densities when compared to older and male subjects. The three oldest male control subjects have negligible SS densities. These observations suggest that reduced SS density is not specific for PD, in agreement with the fact that many conditions such as cognitive function, memory consolidation, pharmacological interventions and pre-PSG conditions have been reported to influence SS density (De Gennaro and Ferrara, 2003; Caporro et al., 2012). Further analysis including more PD and iRBD patients, together with a more in-depth investigation of cognitive decline and disease severity would be needed to evaluate the relation of abnormalities in SS development in the disease process, and the use of SS as a prognostic marker. Additionally, SS density has also been reported decreased for other conditions such as Dementia, Alzheimer's disease (AD) and mild cognitive impairment (Rauchs et al., 2008; Westerberg et al., 2012; Latreille et al., 2015), and is also a sign of normal aging (Wauquier, 1993; De Gennaro and Ferrara, 2003; Ktonas et al., 2009).

To our knowledge, no studies have investigated the impact of L-DOPA on SS morphology. Previous studies have reported that

TABLE 3 | Mean (μ) and standard deviation (σ) for characteristics of spindles in patients with Parkinson's disease (PD) compared to controls (C).

Spindle characteristic	Group consensus (759 SS)		*P*
	PD(-MSA)	*C*	
Duration [sec, $\mu \pm \sigma$]	0.86 ± 0.35	0.77 ± 0.36	<0.001
Frequency [Hz, $\mu \pm \sigma$]	12.27 ± 1.07	12.80 ± 1.23	<0.001
Max peak-to-peak amplitude [μV, $\mu \pm \sigma$]	56.35 ± 18.97	48.19 ± 15.55	<0.001
Max peak-to-peak amplitude After removal of frequencies < 4 Hz [μV, $\mu \pm \sigma$]	53.42 ± 17.84	45.29 ± 14.41	<0.001
Percent-to-peak amplitude [$\mu \pm \sigma$]	0.47 ± 0.23	0.46 ± 0.23	NS
Density [per min, $\mu \pm \sigma$]	0.72 ± 1.28	1.86 ± 1.57	<0.007

In this case, the patient that later was diagnosed with Multiple System Atrophy (MSA) was excluded from the PD group [PD (-MSA)]. P-values for the Wilcoxon rank sum tests between the two groups are shown. Only spindles in the group consensus are included in the comparison.

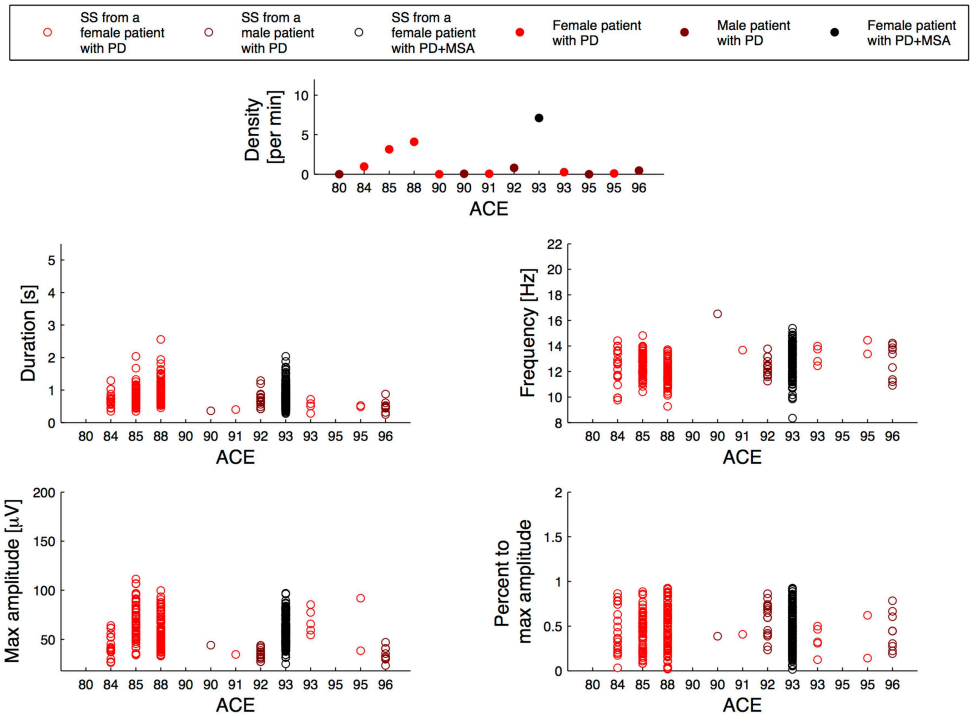


FIGURE 3 | Distribution of the morphology measures for the spindles from 13/15 patients with Parkinson's disease (PD), where the patients are sorted according to their Addenbrooke's cognitive examination (ACE) scores. PD + MSA indicates a patient with PD, that later developed Multiple System Atrophy (MSA).

SS density is increased in patients with PD taking dopaminergic treatment compared to non-treated patients, but the study lacks a comparison to controls, and evaluation of spindle morphology (Puca et al., 1973). As dopaminergic treatments were not discontinued in this study, we cannot rule out that the changes in SS morphology observed are due to the dopaminergic interactions from the treatments, although we do not believe so, as we did not see increases in SS density in these subjects. Future studies will have to investigate this further including a potential association between amount and duration of L-DOPA and/or dopamine agonist treatment and SS morphological changes.

Surprisingly, SS in patients with PD had a longer duration and a higher maximum peak-to-peak amplitude. To our knowledge, no other studies have reported differences in SS duration in patients with PD when compared to controls. The maximum peak-to-peak amplitude significantly differ for SS identifications in the group consensus as well as for each of the individual expert's identifications. This finding was also significant after we

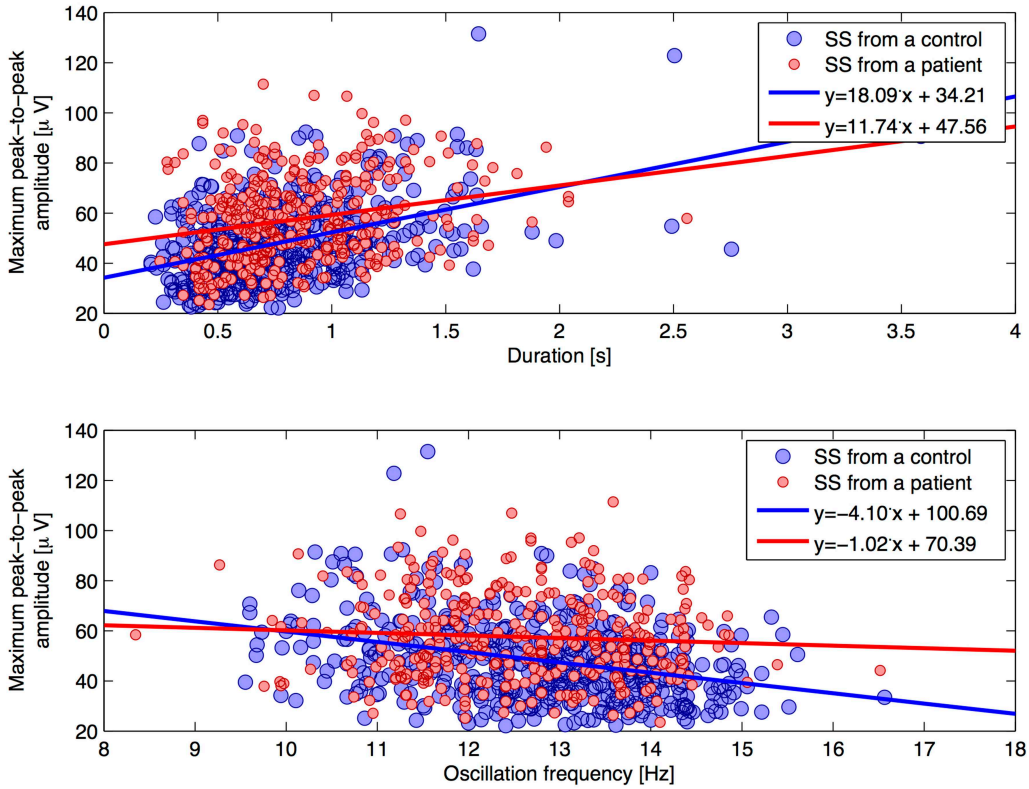


FIGURE 4 | Two scatterplots for individual SS characteristics. The plot illustrates the maximum peak-to-peak amplitude (without removal of frequencies below 4 Hz) as a function of (1) duration (top plot) and (2) oscillation frequency (lower plot), respectively. Trend lines are added for each group.

TABLE 4 | Percent of sleep spindles (SS) identified in the group consensus that do not strictly meet AASM criteria Iber et al. (2007).

AASM criteria	Total SS	PD SS	PD-MSA SS	Control SS	*P*-value PD vs. controls	*P*-value PD-MSA vs. controls
Duration too short (<0.5 s)	0.169	0.128	0.134	0.194	0.010	NS
Duration too long (>3 s)	0.001	0	0	0.002	NS	NS
Oscillation frequency too slow (<11 Hz)	0.097	0.090	0.099	0.101	NS	NS
Oscillation frequency too high (>16 Hz)	0.002	0.003	0.005	0.002	NS	NS
At least one criteria not met	0.253	0.212	0.228	0.278	0.027	NS

There were a total of 344 SS from 11 patients with Parkinson's disease (PD) and 557 SS from 15 control subjects. There were 202 SS from 10 patients when one patient with PD, who later was diagnosed with Multiple System Atrophy (MSA) [PD(-MSA)] was left out. X^2-tests were used to test for significance between spindles from PD patients (including and excluding the one with MSA) and control subjects.

filtered the data to eliminate the impact of low frequency, high amplitude waves. This was surprising, and contradicts the idea that polygraphic features such as SS and K-complexes are less well formed in various NDDs (Petit et al., 2004; Ktonas et al., 2009). By computing maximum peak-to-peak amplitude both without any further filtration and after elimination of low frequencies, our data show that patients with PD show SS with higher amplitudes, regardless of the EEG patterns surrounding them. Margis et al. (2015) reports increased sigma power in N2 sleep of patients with PD vs. controls. Increased sigma power is consistent with our findings of increased duration and amplitude of spindles, which would overpower the decrease in spindle density we and others have reported in PD. Interestingly, SS morphology was unchanged in schizophrenia patients compared to controls, even though they had a significant decrease in SS density (Wamsley et al., 2012).

Enhanced maximum peak-to-peak amplitude is also not consistent with the findings of Latreille et al. (2015), who reports no significant differences of SS amplitude between PD patients and controls, and significantly reduced SS amplitude in patients with PD, who later developed Dementia when compared with controls. The SS in Latreille et al. (2015) were found automatically and mandated a duration criteria of least 0.5 s to be included. Also, the spindle detection method includes a filtration of the signal (11–15 Hz) and a threshold determined based on root-mean-square (RMS) values of the background NREM activity (Martin et al., 2013). Lastly, the SS in Latreille et al. (2015) were detected in all NREM stages, and the individual SS characteristics (amplitude and frequency) were computed as the mean of both hemispheres, as they found no significant hemispheric interaction. The definition of SS is thus not the same in the two studies, and the different results could be due to the fact that automatic detectors detect SS that humans cannot see. Another explanation could be that the detector in Latreille et al. (2015) lack to identify the smaller SS in controls, thereby enlarging the mean spindle amplitude in controls. If the threshold used is based on values across all NREM sleep stages, different amount of NREM stages between controls and patients influences the threshold, maybe resulting in harder thresholds to cross for control spindles. Lastly, taking into account the fact that PD patients show more mixed sleep patterns making sleep stages more difficult to distinguish (Danker-Hopfe et al., 2004; Jensen et al., 2010), it could also be that more N3 sleep is present in the annotated data of patients compared to controls, although we did select data from N2 sleep according to each hypnogram. Whether the contradicting findings are due to methodological reasons only, have to be investigated in future studies, e.g., by applying different automatic spindle detectors on the same dataset and on data from different derivations, and see if the morphological alterations are consistent across detectors, manually scorings and derivations.

EEG slowing has been frequently reported in PD (Petit et al., 2004; Rodrigues Brazète et al., 2013), including slowing in occipital, temporo-occipital and frontal regions (Sirakov and Mezan, 1963; Soikkeli et al., 1991; Primavera and Novello, 1992). It is therefore not surprising that we found slower SS oscillation frequencies in PD patients. Whether or not this is specific for PD or generalizable to other NDDs will need further investigations.

TABLE 5 | Mean (μ) and standard deviation (σ) for the spindle characteristics found for the spindles in the group consensus meeting the AASM criteria.

Spindle characteristic	Group consensus (673 SS)		*P*
	PD	*C*	
Duration [sec, $\mu \pm \sigma$]	0.93 ± 0.33	0.85 ± 0.31	$1.95 \cdot 10^{-4}$
Frequency [Hz, $\mu \pm \sigma$]	12.65 ± 1.01	13.00 ± 0.96	$9.04 \cdot 10^{-6}$
Max peak-to-peak amplitude [μV, $\mu \pm \sigma$]	58.97 ± 16.64	48.60 ± 14.92	$3.90 \cdot 10^{-16}$
Max peak-to-peak amplitude After removal of frequencies < 4 Hz [μV, $\mu \pm \sigma$]	56.06 ± 15.75	45.76 ± 13.89	$5.27 \cdot 10^{-18}$
Percent-to-peak amplitude [$\mu \pm \sigma$]	0.47 ± 0.23	0.45 ± 0.23	NS
Density [per min, $\mu \pm \sigma$]	0.90 ± 1.71	1.34 ± 1.25	$4.50 \cdot 10^{-2}$

Wilcoxon rank sum tests were used to test for significance between patients with PD and control subjects (C).

TABLE 6 | Mean (μ) and standard deviation (σ) for the inter-expert reliability measure F_1-scores and Cohen's Kappa (κ) for scoring sleep spindles (SS).

SS group definition	F_1-score		κ		*P*
	PD	*C*	*PD*	*C*	
Low confidence "maybe"	0.12 ± 0.11	0.17 ± 0.12	0.14 ± 0.11 "slight"	0.16 ± 0.12 "slight"	NS
Medium confidence "probably"	0.13 ± 0.10	0.19 ± 0.11	0.15 ± 0.10 "slight"	0.18 ± 0.11 "slight"	NS
High confidence "definitely"	0.24 ± 0.13	0.32 ± 0.13	0.21 ± 0.13 "fair"	0.32 ± 0.13 "fair"	$4.76 \cdot 10^{-2\kappa}$
Medium or high confidence "probably/definitely"	0.34 ± 0.15	0.39 ± 0.17	0.28 ± 0.15 "fair"	0.39 ± 0.17 "fair"	NS
All SS	0.41 ± 0.16	0.45 ± 0.15	0.32 ± 0.17 "fair"	0.43 ± 0.15 "moderate"	NS

The mean and standard deviations are taken across the ten expert-pairs available. Wilcoxon rank sum tests were used to test for significantly lower inter-expert reliability for scoring SS in patients with Parkinson's disease (PD) compared to control subjects (C). $^{\kappa}$ indicates significance for κ and F indicates significance for F_1-score.

In AD, Rauchs et al. (2008) found no change in spindle density but found that fast spindles (defined as having frequencies of 13–15 Hz) were significantly reduced when compared to age-matched controls. Consistently, Westerberg et al. (2012) found that patients with amnestic mild cognitive impairment had fewer N2 spindles compared to age-matched controls, and that the reduction was seen in fast spindles (13–15 Hz) and not in slow spindles (11–13 Hz). Latreille et al. (2015) found significant lower SS frequency in patients with PD who later developed Dementia compared to controls, but not in Dementia-free patients with PD compared to controls. This last study might however suffer from a selection bias as they automatically defined SS within a certain frequency range, as stated by the AASM. Nonetheless, as in this study, we found that PD patients had a slower SS frequency, both when looking at SS included in the group consensus, but also when looking at SS strictly meeting AASM criteria.

Figures 2, **3** and **Supplementary Figures 1**, **2** report on SS measures for the PD group consensus, but with subjects sorted according to their disease duration (**Figure 2**), their ACE score (**Figure 3**), their H and Y stage (**Supplementary Figure 1**) and UPDRS part III score (**Supplementary Figure 2**). Although no clear tendency was seen for any of the SS measures for disease duration, ACE score, H and Y stage or UPDRS part III score, longitudinal studies are likely needed to determine whether SS morphology measures can provide prognostic value. Indeed, the patients included here may have had a PD diagnosis for various amounts of time, and inter-subject variation of disease progression and severity makes such a relationship very complicated to analyze. ACE is a brief assessment of cognitive functions and is in this study used as a screening tool to determine Dementia, which none of the patients had at inclusion. A more in-depth examination of cognitive functions as well as a follow-up study of the patients is needed to determine the subject-specific progression and severity rate. These rates can be compared to the SS morphology measures to investigate the prognostic value.

A biomarker does not have to be specific to a disease to have clinical utility, and combining the different SS measures may reveal that different diseases show different trends or different combinations of changes in SS morphology measures. If a trend is found, it is important to also look at SS that might fall out of the stated AASM criteria, as not doing that may misrepresent the data. **Table 4** shows that a rather high proportion of SS in both groups do not meet AASM criteria. Additionally, when looking at inter-expert reliability, it was found that experts are less likely to agree on definite SS in patients when compared to controls. Considering that automatic SS detectors are likely to be used in patients with NDDs, it is highly encouraged to build detectors capable of detecting atypical SS as well. Such atypical SS could be spindles with abnormal duration or frequency or spindles surrounded by EEG that is not typically seen in N2 sleep. Because of this, detectors should not be constrained or designed to perform well only in the context of a single expert or for normal EEG. Ideally, automatic detectors should give a confidence score for each detected SS and group subtypes of SS using specific parameters describing their morphology. Specifically, description of "probable SS" in different patient groups may give a better idea of the specific morphological changes that can be observed for each disease. Also, such studies should investigate how disease duration and/or severity impact morphology. Such in-depth studies would be beneficial to better understand the pathological differences between the NDDs and also see if any of the morphology measures hold potential for separating diseases or subtypes of them.

In conclusion, we investigated SS in an objective way and found that the oscillation frequency and duration of SS manually scored in clinical settings are not necessarily bound to the limits given by AASM. The shorter or slower SS must have had an ability to stand out from the background EEG, and we believe that these per-definition-not-SS should be included in studies analyzing SS morphology changes, particularly when searching for disease biomarkers.

Based on a group consensus of five individual experts' identification of SS in N2 sleep, we compared 15 patients with PD with 15 age-matched control subjects and found that patients show a lower SS density and that their SS have a longer duration, a higher maximum peak-to-peak amplitude and a slower oscillation frequency. All the included patients were taking dopaminergic treatment, and we can therefore not rule out that the significant differences found could be due to treatment effects. We conclude that SS are significantly altered in patients with PD, but that due to high inter-subject variability in disease progression and severity, future longitudinal studies are needed to investigate the clinical utility of the SS morphology changes as well as their value as prognostic biomarkers.

Financial Support

The PhD project is supported by grants from H. Lundbeck A/S, the Lundbeck Foundation, the Technical University of Denmark and the Center for Healthy Aging, University of Copenhagen.

Acknowledgments

The authors would like to thank the five experts for their time and effort in annotating and giving confidence scores of the sleep spindles analyzed in this study. The PhD project is supported by grants from H. Lundbeck A/S, the Lundbeck Foundation, the Technical University of Denmark, Center for Healthy Aging, University of Copenhagen and Stanford Center for Sleep Sciences and Medicine.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnhum.2015.00233/abstract

Supplementary Figure 1 | Distribution of the morphology measures for the spindles from 11/15 patients with Parkinson's disease (PD), where the patients are sorted according to their Hoehn and Yahr (H and Y) stage.

Supplementary Figure 2 | Distribution of the morphology measures for the spindles from 11/15 patients with Parkinson's disease (PD), where the patients are sorted according to their Unified Parkinson's Disease Rating Scale (UPDRS) part III score.

References

American Academy of Sleep Medicine. (2014). *International Classification of Sleep Disorders - Third Edition (ICSD-3), 3rd Edn*. Westchester, IL: American Academy of Sleep Medicine.

Braak, H., Tredici, K., Del Rüb, U., Vos, R. A., de Steur, E. N. H. J., and Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* 24, 197–211. doi: 10.1016/S0197-4580(02)00065-9

Caporro, M., Haneef, Z., Yeh, H. J., Lenartowicz, A., Buttinelli, C., Parvizi, J., et al. (2012). Functional MRI of sleep spindles and K-complexes. *Clin. Neurophysiol.* 123, 303–309. doi: 10.1016/j.clinph.2011.06.018

Chaudhuri, K. R., Healy, D. G., and Schapira, A. H. (2006). Non-motor symptoms of Parkinson's disease: diagnosis and management. *Lancet Neurol.* 5, 235–245. doi: 10.1016/S1474-4422(06)70373-8

Chaudhuri, K. R., Odin, P., Antonini, A., and Martinez-Martin, P. (2011). Parkinson's disease: the non-motor issues. *Parkinsonism Relat. Disord.* 17, 717–723. doi: 10.1016/j.parkreldis.2011.02.018

Christensen, J. A. E., Frandsen, R., Kempfner, J., Arvastson, L., Christensen, S. R., Jennum, P., et al. (2012). Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2012, 2941–2944. doi: 10.1109/EMBC.2012.6346580

Christensen, J. A. E., Kempfner, J., Zoetmulder, M., Leonthin, H. L., Arvastson, L., Christensen, S. R., et al. (2014a). Decreased sleep spindle density in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. *Clin. Neurophysiol.* 125, 512–519. doi: 10.1016/j.clinph.2013.08.013

Christensen, J. A. E., Koch, H., Frandsen, R., Kempfner, J., Arvastson, L., Christensen, S. R., et al. (2013). Classification of iRBD and Parkinson's patients based on eye movements during sleep. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2013, 441–444. doi: 10.1109/EMBC.2013.6609531

Christensen, J. A. E., Zoetmulder, M., Koch, H., Frandsen, R., Arvastson, L., Christensen, S. R., et al. (2014b). Data-driven modeling of sleep EEG and EOG reveals characteristics indicative of pre-Parkinson's and Parkinson's disease. *J. Neurosci. Methods* 235, 262–276. doi: 10.1016/j.jneumeth.2014.07.014

Comella, C. L., Tanner, C. M., and Ristanovic, R. K. (1993). Polysomnographic sleep measures in Parkinson's disease patients with treatment-induced hallucinations. *Ann. Neurol.* 34, 710–714. doi: 10.1002/ana.410340514

Danker-Hopfe, H., Kunz, D., Gruber, G., Klösch, G., Lorenzo, J. L., Himanen, S. L., et al. (2004). Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J. Sleep Res.* 13, 63–69. doi: 10.1046/j.1365-2869.2003.00375.x

De Gennaro, L., and Ferrara, M. (2003). Sleep spindles: an overview. *Sleep Med. Rev.* 7, 423–440. doi: 10.1053/smrv.2002.0252

Emser, W., Brenner, M., Stober, T., and Schimrigk, K. (1988). Changes in nocturnal sleep in Huntington's and Parkinson's disease. *J. Neurol.* 235, 177–179. doi: 10.1007/BF00314313

Fogel, S., Martin, N., Lafortune, M., Barakat, M., Debas, K., Laventure, S., et al. (2012). NREM sleep oscillations and brain plasticity in aging. *Front. Neurol.* 3:176. doi: 10.3389/fneur.2012.00176

Fogel, S. M., and Smith, C. T. (2011). The function of the sleep spindle: a physiological index of intelligence and a mechanism for sleep-dependent memory consolidation. *Neurosci. Biobehav. Rev.* 35, 1154–1165. doi: 10.1016/j.neubiorev.2010.12.003

Galvin, J. E., Lee, M.-Y., and Trojanowski, J. Q. (2001). Synucleinopathies: clinical and pathological implications. *Arch. Neurol.* 58, 186–190. doi: 10.1001/archneur.58.2.186

Garcia-Ruiz, P. J., Chaudhuri, K. R., and Martinez-Martin, P. (2014). Non-motor symptoms of Parkinson's disease A review…from the past. *J. Neurol. Sci.* 338, 30–33. doi: 10.1016/j.jns.2014.01.002

Happe, S., Anderer, P., Pirker, W., Klösch, G., Gruber, G., Saletu, B., et al. (2004). Sleep microstructure and neurodegeneration as measured by [123I]beta-CIT SPECT in treated patients with Parkinson's disease. *J. Neurol.* 251, 1465–1471. doi: 10.1007/s00415-004-0564-3

Iber, C., Ancoli-Israel, S., Chesson, A. L., and Quan, S. F. (2007). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Tchnical Specification*. Westchester, IL: American Academy of Sleep Medicine.

Iranzo, A. (2011). Sleep-wake changes in the premotor stage of Parkinson disease. *J. Neurol. Sci.* 310, 283–285. doi: 10.1016/j.jns.2011.07.049

Jensen, P. S., Sorensen, H. B. D., Leonthin, H. L., and Jennum, P. (2010). Automatic sleep scoring in normals and in individuals with neurodegenerative disorders according to new international sleep scoring criteria. *J. Clin. Neurophysiol.* 27, 296–302. doi: 10.1097/WNP.0b013e3181eaad4b

Kempfner, J., Jennum, P., Nikolic, M., Christensen, J. A. E., and Sorensen, H. B. D. (2013). Sleep phenomena as an early biomarker for Parkinsonism. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2013, 5773–5776. doi: 10.1109/EMBC.2013.6610863

Koch, H., Christensen, J. A. E., Frandsen, R., Arvastson, L., Christensen, S. R., Jennum, P. J., et al. (2014). Automatic sleep classification using a data-driven topic model reveals latent sleep states. *J. Neurosci. Methods* 235, 130–137. doi: 10.1016/j.jneumeth.2014.07.002

Ktonas, P. Y., Golemati, S., Xanthopoulos, P., Sakkalis, V., Ortigueira, M. D., Tsekou, H., et al. (2009). Time-frequency analysis methods to quantify the time-varying microstructure of sleep EEG spindles: possibility for dementia biomarkers? *J. Neurosci. Methods* 185, 133–142. doi: 10.1016/j.jneumeth.2009.09.001

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi: 10.2307/2529310

Latreille, V., Carrier, J., Lafortune, M., Postuma, R. B., Bertrand, J.-A., Panisset, M., et al. (2015). Sleep spindles in Parkinson's disease may predict the development of dementia. *Neurobiol. Aging* 36, 1083–1090. doi: 10.1016/j.neurobiolaging.2014.09.009

Latreille, V., Carrier, J., Montplaisir, J., Lafortune, M., and Gagnon, J.-F. (2011). Non-rapid eye movement sleep characteristics in idiopathic REM sleep behavior disorder. *J. Neurol. Sci.* 310, 159–162. doi: 10.1016/j.jns.2011.06.022

Leiser, S. C., Dunlop, J., Bowlby, M. R., and Devilbiss, D. M. (2011). Aligning strategies for using EEG as a surrogate biomarker: a review of preclinical and clinical research. *Biochem. Pharmacol.* 81, 1408–1421. doi: 10.1016/j.bcp.2010.10.002

Margis, R., Schönwald, S. V., Carvalho, D. Z., Gerhardt, G. J. L., and Rieder, C. R. M. (2015). NREM sleep alpha and sigma activity in Parkinson's disease: evidence for conflicting electrophysiological activity? *Clin. Neurophysiol.* 126, 951–958. doi: 10.1016/j.clinph.2014.07.034

Martin, N., Lafortune, M., Godbout, J., Barakat, M., Robillard, R., Poirier, G., et al. (2013). Topography of age-related changes in sleep spindles. *Neurobiol. Aging* 34, 468–476. doi: 10.1016/j.neurobiolaging.2012.05.020

Micallef, J., Rey, M., Eusebio, A., Audebert, C., Rouby, F., Jouve, E., et al. (2009). Antiparkinsonian drug-induced sleepiness: a double-blind placebo-controlled study of L-dopa, bromocriptine and pramipexole in healthy subjects. *Br. J. Clin. Pharmacol.* 67, 333–340. doi: 10.1111/j.1365-2125.2008.03310.x

Micanovic, C., and Pal, S. (2014). The diagnostic utility of EEG in early-onset dementia: a systematic review of the literature with narrative analysis. *J. Neural Transm.* 121, 59–69. doi: 10.1007/s00702-013-1070-5

Myslobodsky, M., Mintz, M., Ben-Mayor, V., and Radwan, H. (1982). Unilateral dopamine deficit and lateral EEG assymmetry: sleep abnormalities in hemi-Parkinson's patients. *Electroencephalogr. Clin. Neurophysiol.* 54, 227–231. doi: 10.1016/0013-4694(82)90164-X

Nguyen, L., Bradshaw, J. L., Stout, J. C., Croft, R. J., and Georgiou-Karistianis, N. (2010). Electrophysiological measures as potential biomarkers in Huntington's disease: review and future directions. *Brain Res. Rev.* 64, 177–194. doi: 10.1016/j.brainresrev.2010.03.004

Nicolas, A., Petit, D., Rompré, S., and Montplaisir, J. (2001). Sleep spindle characteristics in healthy subjects of different age groups. *Clin. Neurophysiol.* 112, 521–527. doi: 10.1016/S1388-2457(00)00556-3

Petit, D., Gagnon, J.-F., Fantini, M. L., Ferini-Strambi, L., and Montplaisir, J. (2004). Sleep and quantitative EEG in neurodegenerative disorders. *J. Psychosom. Res.* 56, 487–496. doi: 10.1016/j.jpsychores.2004.02.001

Postuma, R. B., Gagnon, J. F., Rompré, S., and Montplaisir, J. Y. (2010). Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease. *Neurology* 74, 239–244. doi: 10.1212/WNL.0b013e3181ca0166

Primavera, A., and Novello, P. (1992). Quantitative electroencephalography in Parkinson's disease, dementia, depression and normal aging. *Neuropsychobiology* 25, 102–105. doi: 10.1159/000118817

Puca, F., Bricolo, A., and Turella, G. (1973). Effect of L-DOPA or amantadine therapy on sleep spindles in parkinsonism. *Electroencephalogr. Clin. Neurophysiol.* 35, 327–330. doi: 10.1016/0013-4694(73)90245-9

Rauchs, G., Schabus, M., Parapatics, S., Bertran, F., Clochon, P., Hot, P., et al. (2008). Is there a link between sleep changes and memory in Alzheimer's disease? *Neuroreport* 19, 1159–1162. doi: 10.1097/WNR.0b013e32830867c4

Rodrigues Brazète, J., Montplaisir, J., Petit, D., Postuma, R. B., Bertrand, J.-A., Génier Marchand, D., et al. (2013). Electroencephalogram slowing in rapid eye movement sleep behavior disorder is associated with mild cognitive impairment. *Sleep Med.* 14, 1059–1063. doi: 10.1016/j.sleep.2013.06.013

Roth, C., Jeanmonod, D., Magnin, M., Morel, A., and Achermann, P. (2000). Effects of medial thalamotomy and pallido-thalamic tractotomy on sleep and waking EEG in pain and parkinsonian patients. *Clin. Neurophysiol.* 111, 1266–1275. doi: 10.1016/S1388-2457(00)00295-9

Salawu, F., Danburam, A., and Olokoba, A. (2010). Non-motor symptoms of Parkinson's disease: diagnosis and management. *Niger. J. Med.* 19, 126–131. doi: 10.4314/njm.v19i2.56496

Schabus, M., Hödlmoser, K., Gruber, G., Sauter, C., Anderer, P., Klösch, G., et al. (2006). Sleep spindle-related activity in the human EEG and its relation to general cognitive and learning abilities. *Eur. J. Neurosci.* 23, 1738–1746. doi: 10.1111/j.1460-9568.2006.04694.x

Schenck, C. H., Boeve, B. F., and Mahowald, M. W. (2013a). Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series. *Sleep Med.* 14, 744–748. doi: 10.1016/j.sleep.2012.10.009

Schenck, C. H., Bundlie, S. R., and Mahowald, M. W. (1996). Delayed emergence of a parkinsonian disorder in 38 % of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder. *Neurology* 46, 388–393. doi: 10.1212/WNL.46.2.388

Schenck, C. H., Montplaisir, J. Y., Frauscher, B., Hogl, B., Gagnon, J.-F., Postuma, R., et al. (2013b). Rapid eye movement sleep behavior disorder: devising controlled active treatment studies for symptomatic and neuroprotective therapy–a consensus statement from the International Rapid Eye Movement Sleep Behavior Disorder Study Group. *Sleep Med.* 14, 795–806. doi: 10.1016/j.sleep.2013.02.016

Sirakov, A., and Mezan, I. (1963). EEG findings in parkinsonism. *Electroencephalogr. Clin. Neurophysiol.* 15, 321–322. doi: 10.1016/0013-4694(63)90101-9

Soikkeli, R., Partanen, J., Soininen, H., Pääkkönen, A., and Sr. Riekkinen, P. (1991). Slowing of EEG in Parkinson's disease. *Electroencephalogr. Clin. Neurophysiol.* 79, 159–165. doi: 10.1016/0013-4694(91)90134-P

Sorensen, G. L., Kempfner, J., Zoetmulder, M., Sorensen, H. B. D., and Jennum, P. (2012). Attenuated heart rate response in REM sleep behavior disorder and Parkinson's disease. *Mov. Disord.* 27, 888–894. doi: 10.1002/mds.25012

Sorensen, G. L., Mehlsen, J., and Jennum, P. (2013). Reduced sympathetic activity in idiopathic rapid-eye-movement sleep behavior disorder andParkinson's disease. *Auton. Neurosci.* 179, 138–141. doi: 10.1016/j.autneu.2013.08.067

Steriade, M., McCormick, D., and Sejnowski, T. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262, 679–685. doi: 10.1126/science.8235588

Steriade, M., and Timofeev, I. (2003). Neuronal plasticity in thalamocortical networks during sleep and waking oscillations. *Neuron* 37, 563–576. doi: 10.1016/S0896-6273(03)00065-5

Stevens, S., and Comella, C. L. (2013). "Rapid eye movement sleep behavior disorder," in *Parkinson's Disease and Nonmotor Dysfunction,* 2nd Edn., eds R. F. Pfeiffer and Bodis-Wollner (New York, NY: Humana Press Inc.), 257–265.

Wamsley, E. J., Tucker, M. A., Shinn, A. K., Ono, K. E., McKinley, S. K., Ely, A. V., et al. (2012). Reduced sleep spindles and spindle coherence in schizophrenia: mechanisms of impaired memory consolidation? *Biol. Psychiatry* 71, 154–161. doi: 10.1016/j.biopsych.2011.08.008

Warby, S. C., Wendt, S. L., Welinder, P., Munk, E. G. S., Carrillo, O., Sorensen, H. B. D., et al. (2014). Sleep-spindle detection: crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nat. Methods* 11, 385–392. doi: 10.1038/nmeth.2855

Wauquier, A. (1993). Aging and changes in phasic events during sleep. *Physiol. Behav.* 54, 803–806. doi: 10.1016/0031-9384(93)90095-W

Wendt, S. L., Welinder, P., Sorensen, H. B. D., Jennum, P. J., Peppard, P. E., Perona, P., et al. (2014). Inter-expert and intra-expert reliability in sleep spindle scoring. *Clin. Neurophysiol.* doi: 10.1016/j.clinph.2014.10.158. [Epub ahead of print].

Wermuth, L., Lassen, C. F., Himmerslev, L., Olsen, J., and Ritz, B. (2012). Validation of hospital register-based diagnosis of Parkinson's disease. *Dan. Med. J.* 59:A4391.

Westerberg, C. E., Mander, B. A., Florczak, S. M., Weintraub, S., Mesulam, M.-M., Zee, P. C., et al. (2012). Concurrent impairments in sleep and memory in amnestic mild cognitive impairment. *J. Int. Neuropsychol. Soc.* 18, 490–500. doi: 10.1017/S135561771200001X

Conflict of Interest Statement: The Reviewer Veronique Latreille declares that, despite being affiliated to the same institution as the author Simon C. Warby, the review process was handled objectively and no conflict of interest exists. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Patent I

TITLE:
Method for detection of an abnormal sleep pattern in a person

INVENTORS:
Julie A. E. Christensen, Helge B. D. Sorensen, Poul Jennum, Søren R. Christensen and Lars Arvastson

APPLICATION NO.:
PCT/EP2013/067297

PUBLICATION NO.:
WO 2014/029764

APPLICANT:
Technical University of Denmark

DATE OF PRIORITY:
20 August 2012

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 February 2014 (27.02.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/029764 A1

(51) International Patent Classification:
A61B 5/0496 (2006.01) *A61B 5/00* (2006.01)

(21) International Application Number:
PCT/EP2013/067297

(22) International Filing Date:
20 August 2013 (20.08.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12181048.5 20 August 2012 (20.08.2012) EP
13174352.8 28 June 2013 (28.06.2013) EP

(71) Applicant: **DANMARKS TEKNISKE UNIVERSITET** [DK/DK]; Anker Engelundsvej 1, DK-2800 Kgs. Lyngby (DK).

(72) Inventors: **SØRENSEN, Helge Bjarup Dissing**; Baunevej 2A, DK-3230 Græsted (DK). **CHRISTENSEN, Julie Anja Engelhard**; Gothersgade 10C, 2.tv, DK-1123 Copenhagen C (DK). **JENNUM, Poul Jørgen**; Skovbakken 100, DK-3520 Farum (DK). **CHRISTENSEN, Søren Rahn**; Egernbo 9, DK-2665 Vallensbæk Strand (DK). **ARVASTSON, Lars**; Roslins Väg 18C, S-217 55 Malmö (SE).

(74) Agent: **HØIBERG A/S**; St. Kongensgade 59 A, DK-1264 Copenhagen K (DK).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

[Continued on next page]

(54) Title: METHOD FOR DETECTION OF AN ABNORMAL SLEEP PATTERN IN A PERSON

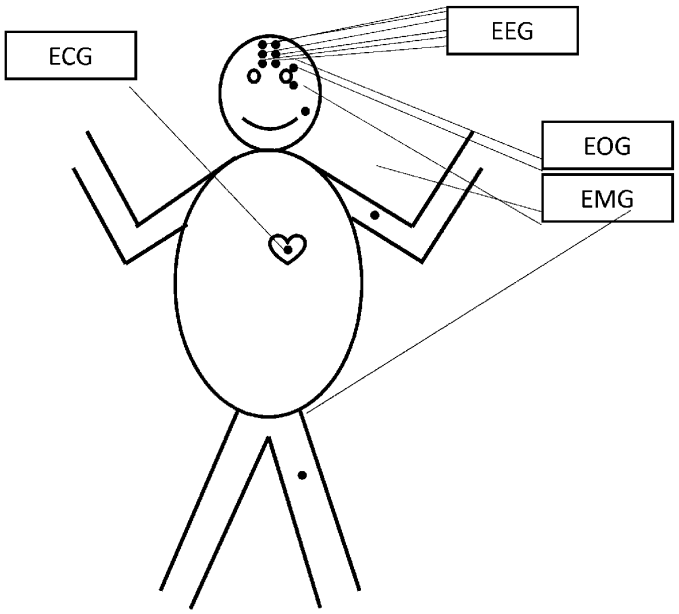


Fig. 5

(57) Abstract: The present disclosure relates to a method for detection of an abnormal sleep pattern based on a dataset of Electrooculography (EOG) signals obtained from a sleeping subject over a time interval, the method comprising the steps of dividing the time interval into a plurality of subintervals, each subinterval preferably corresponding to a sleep epoch, classifying each subinterval in terms of sleep stages, thereby obtaining a temporal sleep stage pattern, wherein a subject having an uncharacteristic temporal distribution of sleep stages is characterized as having an abnormal sleep pattern.

WO 2014/029764 A1

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

Patent II

TITLE:
Sleep spindles as biomarker for early detection of neurodegenerative disorders

INVENTORS:
Julie A. E. Christensen, Lykke Kempfner, Poul Jennum, Helge B. D. Sorensen, Lars Arvastson and Søren R. Christensen

APPLICATION NO.:
14/290,402

PUBLICATION NO.:
US 2015/0080671

APPLICANT:
Technical University of Denmark

DATE OF PRIORITY:
29 May 2013

US 20150080671A1

(19) **United States**

(12) **Patent Application Publication**

Christensen et al.

(10) **Pub. No.: US 2015/0080671 A1**

(43) **Pub. Date: Mar. 19, 2015**

(54) **SLEEP SPINDLES AS BIOMARKER FOR EARLY DETECTION OF NEURODEGENERATIVE DISORDERS**

(71) Applicants: **Technical University of Denmark**, Kgs. Lyngby (DK); **Glostrup Hospital**, Glostrup (DK); **H. Lundbeck A/S**, Valby (DK)

(72) Inventors: **Julie Anja Engelhard Christensen**, Copenhagen K (DK); **Lykke Kempfner**, Herlev (DK); **Poul Jørgen Jennum**, Farum (DK); **Helge Bjarup Dissing Sørensen**, Graested (DK); **Lars Arvastson**, Malmo (SE); **Søren Rahn Christensen**, Vallensbaek Strand (DK)

(21) Appl. No.: **14/290,402**

(22) Filed: **May 29, 2014**

(30) **Foreign Application Priority Data**

May 29, 2013 (EP) 13169679.1

Publication Classification

(51) **Int. Cl.**

A61B 5/00	(2006.01)
A61B 5/11	(2006.01)
A61B 5/0496	(2006.01)
A61B 5/0476	(2006.01)

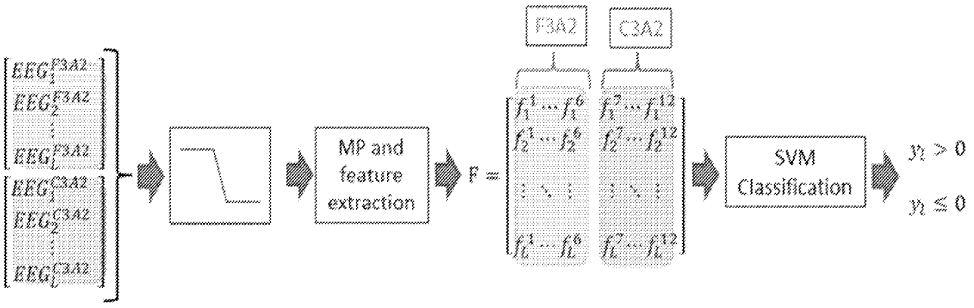
(52) **U.S. Cl.**

CPC ***A61B 5/4812*** (2013.01); ***A61B 5/0476*** (2013.01); ***A61B 5/11*** (2013.01); ***A61B 5/0496*** (2013.01); ***A61B 5/4082*** (2013.01)

USPC **600/301**; 600/544

(57) **ABSTRACT**

The present invention relates to the use of sleep spindles as a novel biomarker for early diagnosis of synucleinopathies, in particular Parkinson's disease (PD). The method is based on automatic detection of sleep spindles. The method may be combined with measurements of one or more further biomarkers derived from polysomnographic recordings.



www.elektro.dtu.dk
Department of Electrical Engineering
Biomedical Engineering
Technical University of Denmark
Ørsteds Plads
Building 348
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 38 00
Fax: (+45) 45 93 16 34
Email: info@elektro.dtu.dk